PARTICULATE MODELING AND

CONTROL STRATEGY FOR ATLANTA, GEORGIA

A Dissertation
Presented to
The Academic Faculty

by

Sun-Kyoung Park

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
December 2005

PARTICULATE MODELING AND

CONTROL STRATEGY FOR ATLANTA, GEORGIA

Approved by:

Dr. Armistead G. Russell, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Michael H. Bergin
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. James A. Mulholland
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Athanasios Nenes
School of Earth and Atmospheric Sciences / School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Yuhang Wang
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Date Approved: October 2005

To my parents and husband

ACKNOWLEDGEMENTS

I would like to thank a number of individuals for their assistance and support to make this dissertation possible. Foremost, I would like to express my sincerest gratitude to my advisor, Dr. Armistead Russell. His insight and knowledge in the field of air quality modeling with his hearty support made my whole program of study really fruitful. When I should be out of town due to my personal problem, he also provided me an office remotely, and supported me so that I could continue my research. If he did not encourage me at that time, I might not continue my study any more. His consideration and encouragement would remain in my mind for my life. I am also thankful to my committee members, Dr. Michael Bergin, Dr. James Mulholland, Dr. Yuhang Wang, and Dr. Athanasios Nenes for their valuable time, constructive criticisms, and suggestions. I enjoyed many interesting and fruitful discussion we had on my research work.

I also thank my friends, faculty, and staff at the School of Civil and Environmental Engineering program too many to name all of them. Special thanks go to group members, Yongtao Hu, Di Tian, Jaemeen Baek, Sergey Napelenok, Alper Unal, Kasemsan Manomaiphiboon, Amit Marmur, Daniel Cohan, Sangil Lee, K.J. Liao, Evan Cobb, Burcak Kaynak, Santosh Chandru, Michelle Bergin, and Talat Odman. They made my life at Georgia Tech more enjoyable and fruitful. Finally, I have deep appreciation to my parents, sister, and brother for their unconditional love and concerns. Also, I want to thank my daughter, Janice Kim, for growing healthily. Above all, I am blessed by the love and encouragement of my husband, Seoung Bum Kim. Completing the Ph.D. would have been impossible without his support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Ambient fine particles reduce visibility in urban regions, deposit acid, and disturb the Earth's radiation balance. Elevated particle levels also adversely affect human health. As a response to an increasing number of scientific studies that link high amounts of fine particles and adverse health effects, in 1997, by the US EPA, National Ambient Air Quality Standards (NAAQSs) for $PM_{2.5}$ were promulgated. $PM_{2.5}$ levels in the metropolitan Atlanta, Georgia area exceed the NAAQS. Thus, control strategies are needed to reduce fine particle levels. This thesis develops the control strategies of $PM_{2.5}$ in the Atlanta area, and investigates the various uncertainties residing in air quality models. In addition, this study proposes strategies to ameliorate the accuracy of the air quality models.

$PM_{2.5}$ levels have been measured in the metropolitan Atlanta, Georgia area since 1999. Based on the measured $PM_{2.5}$ concentrations, an emission control level meeting the NAAQS for $PM_{2.5}$ is estimated. The property that $PM_{2.5}$ data fit well to the lognormal distribution is used to account for the level of uncertainty in calculating the required controls. Results show that the needed emission reductions to meet the NAAQS in the Atlanta area range from 18% to 26% at 50% confidence interval (CI), and range from 22% to 40% at 95% CI, depending on the location.

The estimated control levels can be tested using air quality models to better assess if the proposed levels will achieve sufficient reduction in $PM_{2.5}$. For the models to be used in such application with confidence, the model needs to be evaluated. Numerous studies show that the model evaluation itself has uncertainty due to the evaluation

method. Uncertainties caused by spatial variations of pollutant concentrations were analyzed. Eulerian three-dimensional photochemical models simulate volume-averaged concentrations. Because the volume-averaged measurements are seldom available, the evaluation is done by the comparison between the volume-averaged simulated concentrations and point measurements. Thus, spatial variations of pollutant concentrations result in uncertainty in model performance evaluation. One way to quantify the uncertainty introduced is by considering the error between point measurements and interpolated measured concentrations. This is done considering two months, July 2001 and January 2002, for the United States. Results show that the uncertainty levels are comparable to model error, suggesting that the spatial variability should be taken into account when the model evaluation results are interpreted. When only spatially representative stations are used to evaluate the model, introduced uncertainty decreases and the model performance improves.

Despite differences between what an air quality model simulates and how observations are made, the Community Multiscale Air Quality (CMAQ) model simulates observed particle mass concentrations reasonably well. However, CMAQ does not accurately simulate particle number concentrations. The evaluation done using $PM_{2.5}$ number concentrations in Atlanta, Georgia from 1/1/1999 to 8/31/2000 shows that CMAQ significantly underestimates particle number concentrations. Assumed particle size of the primary emission in CMAQ cause number concentrations to be significantly underestimated, while particle density has a small impact. Simulating particle size distributions by three lognormal modes cannot accurately simulate particle number concentrations for ultra-fine particles, particularly during nucleation events. Though, the

use of the Expectation-Maximization (EM) algorithm to estimate size distribution parameters of measured particles suggests that assumed parameters of three lognormal modes in CMAQ are generally reasonable.

One of the applications of CMAQ pursued here is source apportionment of $PM_{2.5}$, which has been commonly performed using the receptor models, such as the Chemical Mass Balance with Molecular Markers (CMB-MM). A comparison is done between $PM_{2.5}$ mass apportioned using CMB-MM and that using CMAQ, and an analysis of the strengths and limitations of each model is performed. $PM_{2.5}$ source apportionment results from the two methods differ due to different spatial scales, fundamental limitations in each model, and uncertainties in the two approaches. Both models have strengths and limitations, and each model's strength can be used to compensate for the other model's limitations.

The differences in the source apportionment results between CMB and CMAQ was partly due to the inaccurate simulation of $PM_{2.5}$ from CMAQ. Accuracy of CMAQ's simulated concentrations relies on accurate inputs, as well as accurate processing. Studies suggest that the emission input is one of the largest uncertainty sources. Four-Dimensional Data Assimilation (FDDA) has been performed to evaluate the emission strengths over the United States in July 2001 and in January 2002. Emission sources evaluated include CO (total), NH_3 (total), SO_2 (total), NO_x (area/mobile/nonroad), NO_x (point), VOC (area/mobile/nonroad/point), VOC (biogenic), POA (total), PEC (total), and PMFINE (total). Results show that base level emissions for CO and SO_2 sources are relatively accurate, whereas NH_3, NO_x, PEC and PMFINE emissions are overestimated. The emission adjustment for POA and VOC emissions is significantly different among

regions. The difference of the emission adjustment between weekdays and weekends is negligible in most cases. The modeling results with adjusted emissions have smaller mean fractional bias than those with base level emissions, but have similar mean fractional error. This result suggests that the FDDA method is a useful tool to reduce the systematic bias in emissions. However, other factors contributing to the model error exist, and need to be corrected to improve the accuracy of the modeling results.

CHAPTER 1

INTRODUCTION

Elevated ambient particle levels in the atmosphere adversely affect human health (Donaldson et al., 2001; Goldberg et al., 2003; Iwai et al., 2005; Ulrich et al., 2002). These particles also reduce visibility in urban regions, deposit acid, and disturb the Earth's radiation balance (Latha and Badarinath, 2005; Satheesh and Moorthy, 2005; Takemura et al., 2005). In 1997, the National Ambient Air Quality Standard (NAAQS) for $PM_{2.5}$ (particles less than 2.5 μm in diameter) was promulgated as a response to an increasing number of scientific studies that linked high fine particle levels and adverse health effects (Brauer et al., 2001; Dockery and Pope, 1994; Dockery et al., 1993; Ebelt et al., 2000; Gold et al., 1999; Pope et al., 1995; Pope et al., 1999; Saldiva et al., 1995; Vedal, 1997). The long-term standard established by NAAQS set the three-year average of the annual mean $PM_{2.5}$ mass concentrations at less than or equal to 15 $\mu g/m^3$; the short-term standard set the three-year average of the 98^{th} percentile of the 24-hour $PM_{2.5}$ mass concentrations at less than or equal to 65 $\mu g/m^3$ (US-EPA, 2004a). In areas that fail to meet this standard, a control strategy is to be developed that will bring levels within the stated limits.

The metropolitan Atlanta, Georgia, area is one such area where $PM_{2.5}$ levels exceed the NAAQS standard. As part of a study to characterize particulate matter in Atlanta, $PM_{2.5}$ concentrations have been measured since 1999 in four locations (Butler et al., 2003): Fort McPherson (FTM) army center, South DeKalb (SDK), Tucker (TUC), and Jefferson Street (JST). The average $PM_{2.5}$ concentrations were 19.7 $\mu g\ m^{-3}$ (JST),

19.0 μg m^{-3} (FTM), 18.2 μg m^{-3} (SDK), and 20.1 μg m^{-3} (TUC) from 1999 to 2003, which are higher than the long-term air quality standard. The required emission reduction to meet the NAAQS has been calculated using the roll back method, and the level of uncertainty was calculated using the distributional property of $PM_{2.5}$ mass concentrations (Georgopoulos and Seinfeld, 1982; Hadley and Toumi, 2003; Kruize et al., 2003; Larsen, 1969; Lu, 2002; Lu and Fang, 2002). Here, distributional properties of $PM_{2.5}$ will be also used to estimate control needs when temporal variations of $PM_{2.5}$ are taken into account. In addition, the influence of the correlation among $PM_{2.5}$ species on control levels will be investigated using the joint probability function of correlated data.

Another method to develop and assess control strategies is the use of three-dimensional photochemical air quality models, which are essential tools for use in air quality management and scientific investigation. However, thorough evaluation of such models is necessary for them to be used with confidence. Traditionally, air quality models were evaluated by the relative agreement between observed and simulated concentrations despite the fact that models predict volume-average concentrations, whereas monitors measure concentrations at a single point in space. These different spatial scales introduce uncertainty into the calculated model performance if pollutant concentrations are spatially inhomogeneous (McNair et al., 1996). Studies have shown that the pollutant concentrations can differ locally depending on the characteristics of the sites (Blanchard et al., 1999; Chan and Hwang, 1996; Goswami et al., 2002; Larson et al., 1989; Restrepo et al., 2004; Sheppard et al., 2001). These studies suggest that point observations can be markedly different from the volume-averaged levels. Therefore, model performance evaluated by comparing the point observation and the volume-averaged prediction may

not represent how well the model actually simulates particulate matter dynamics. The uncertainty in the model evaluation due to the spatial variability of $PM_{2.5}$ will be investigated using a data withholding method. In addition, a way to reduce` the uncertainty in model performance evaluation will be proposed.

One of the comprehensive Eulerian photochemical air quality models is the Community Multiscale Air Quality (CMAQ) modeling system (Byun and Ching, 1999). CMAQ modeling system is comprised of the main CMAQ Chemical Transport Model (CCTM), meteorology-chemistry Interface Processor (MCIP), Initial/Boundary Condition Processor (ICON/BCON), and Photolytic Rate Constant Processor (JPROC). JPROC is base on the RADM (Chang et al., 1987), and computes species specific photolysis rates for a set of predefined zenith angles, latitude, and altitude.

CMAQ system simulates tropospheric ozone, acid deposition, visibility, and particulate matter ($PM_{2.5}$ and PM_{10}) simultaneously by adapting a one-atmosphere perspective (Dennis, 1998), which emphasizes interactions at different dynamic scales and among multi-pollutants. CCTM simulates pollutant concentrations by solving the governing equation with the given meteorology and emissions, which are modeled by Mesoscale meteorological Model version 5 (MM5) (PSU/NCAR, 2003) and Spares Matrix Operator Kernel Emissions (SMOKE) (US-EPA, 2004e), respectively. The governing species conservation equation is as follows:

$$\frac{\partial c_i}{\partial t} + \nabla \cdot (U\ c_i) = \nabla \rho\ D_i \nabla \left(\frac{c_i}{\rho} \right) + R_i (c_1, c_2, ..., c_n, T, t) + S_i (x, t) \qquad \text{i=1, 2, 3, …, n}$$

where, c_i is the concentration of species i, U is wind velocity vector, D_i is molecular diffusivity of species i, R_i is rate of concentration change of species i by chemical

reaction, S_i *(x,t)* is source/sink of *i* at location *x* and time *t*, ρ is air density, and n is the number of predicted species. The conservation equation describes the formation, transport, and fate of air pollutants, including components for processing emissions, meteorology, topography, air quality observations, and chemistry (Russell and Dennis, 2000).

CMAQ uses the governing equation derived in a generalized coordinate system, in which switching from one coordinate to another is easy. Thus, the use of generalized coordinate system allows the use of generic coordinates for the specific science processes within a model. The flexible switching of the coordinates is important because although the model's overall structure is determined by the choice of a coordinate system, the individual science modules can still use their own generic coordinates that best suit the physical processes they model.

CMAQ system is structured to accommodate many different science process modules, each of which includes the critical science processes. The modularity in CMAQ system provides a distinction for different processes through which developments in specific research areas are readily be incorporated. The coarsest level of modularity is the distinction between the system framework and science models. The second level is the division of science sub-models (MM5/SMOKE/CMAQ). The third level of modularity involves a driver module, processor module, data provider module, and a utility module in CCTM. Currently, nine science process classes defined in CCTM are DRIVER (data flow and synchronizes fractional time steps), HADV/VADV (horizontal/vertical advections), ADJCON (mixing ratio conservation property of advection processes), HDIFF/VDIFF (horizontal diffusion/vertical diffusion and deposition), CHEM (gas-

phase chemical reactions), CLOUD (aqueous-phase reactions and cloud mixing), AERO (aerosol dynamics and size distributions), and PING (plume chemistry). Here, gas-phase chemical mechanism and aerosol dynamics are described in detail.

The most popular gas-phase chemical mechanisms used in CMAQ are CB4 (Carter, 1996; Gery et al., 1989), RADM2 (Chang et al., 1987; Stockwell et al., 1990), and SAPRC-99 (Carter et al., 1997). The CB4 mechanism includes 36 species (27 inorganic and 9 organic species) and 93 reactions, of which 11 photolytic reactions. CB4 uses lumped primary organic species except ethene (ETH), isoprene (ISOP), and formaldehyde (FORM). The carbon-bond types include carbon atoms that contain only single bonds (PAR), double-bonded carbon atoms (OLE), 7-carbon ring structures represented by toluene (TOL), 8-carbon ring structures represented by xylene (XYL), the carbonyl group and adjacent carbon atom in acetaldehyde and higher molecular weight aldehydes represented by acetaldehyde (ADL2), and non-reactive carbon atoms (NR).

The RADM2 mechanism implemented in CMAQ contains 57 species and 158 reactions, of which 21 are photolytic. Primary organics are represented by 15 mechanism species, five of which are explicit because of their high emission rates or because of special reactivitiy considerations (methane, ethane, ethane, isoprene, and formaldehyde). The other ten represent groups of organic compounds aggregated on the basis of their reactivity with the hydroxyl radical (OH) and/or their molecular weights. To account for varying reactivity of the different organics that are lumped into a single group, emissions of each organic within a group are weighted by a reactivity factor (F). RADM2 described in Stockwell et al. (1990) has a minor change when implemented in CMAQ. The negative stoichiometry is removed, and methane is assumed to be constant at $4.5*10^{13}$

molecules/cm^{-3}. The SAPRC-99 mechanism uses product yield coefficients and rate constants for over 100 individual organic compounds. Due to computational constraints, individual organic compounds are lumped into groups with corresponding rate constants and product yield coefficients that have been weighted by mole fractions of the individual organics.

CMAQ predict particle concentrations as well as gas-phase species concentrations. Predicting particle concentrations is more challenging than predicting the gas-phase species concentrations because it requires additional information of the particle size distribution. Depending on the particle size, the physical behavior (e.g., deposition velocity) varies. Particles can be directly emitted into air, or also made in the atmosphere from chemical reaction. CMAQ divides directly emitted particles into two groups, fine ($PM_{2.5}$: particles with size less than 2.5 μm) and coarse particles (PM_{10}: particles with size between 2.5 and 10 μm). $PM_{2.5}$ is treated by two interacting sub-distributions (or modes). Aitken (smaller one) mode represents fresh particles either from nucleation or from direct emissions, while accumulation (larger one) mode represents aged particles. The two modes interact with each other through coagulation, and particles in Aitken mode may grow into accumulation mode. PM_{10} is treated by a coarse mode originated from the direct emissions by natural processes such as wind blowing dust, or sea salt particles by wind driven waves on the sea surface. Chemical species of $PM_{2.5}$ treated are fine species sulfate, nitrate, ammonium, water, anthropogenic and biogenic organic carbon, elemental carbon, and other unspecified material of anthropogenic origin. PM_{10} includes sea salt, wind-blown dust, and other unspecified material of anthropogenic origin.

Particles are made in the atmosphere not only from the direct emissions but alao from the nucleation of sulfuric acid in the presence of water vapor. CMAQ calculates particle nucleation rate, J [m^{-3} sec^{-1}], based on the classical binary homogeneous nucleation theory (Wilemski, 1984) incorporating the hydration effect (Kulmala et al., 1998). The production rate of new particle number, J, is calculated from the nucleation of sulfuric acid vapor, assuming the new particle diameter is 2.0 nm.

$$J = \exp(\chi),$$

where,

$$\chi = 25.1289\,N_{sulf} - 4890.8\frac{N_{sulf}}{T} - \frac{1743.3}{T} - 2.2479\,\delta\,N_{sulf}\,RH + 7643.4\frac{X_{al}}{T} - 1.9712\frac{X_{al}\,\delta}{RH}$$

$$N_{sulf} = \ln(\frac{N_{av}}{N_{ac}}),$$

$$\delta = 1 + \frac{T - 273.15}{273.15},$$

$$X_{al} = 1.2233 - \frac{0.0154\,RA}{RA + RH} + 0.0102\ln N_{av} - 0.0415\ln N_{wv} + 0.0016T$$

N_{av}, N_{ac}, and N_{wv} are sulfuric acid vapor concentration [cm^{-3}], critical sulfuric acid concentration [cm^{-3}], and water vapor concentration [cm^{-3}], respectively. RH and RA denote relative humidity and relative acidity divided by 100%. X_{al} is mole fraction of critical nucleus, and T is the temperature in Kelvin. Production rates of new particle number, surface area, and mass are expressed as:

$$\frac{d\,(Number)}{dt} = J$$

$$\frac{d\,(Surface\ Area)}{dt} = \pi d_2^2\, J$$

$$\frac{d\,(Mass)}{dt} = \frac{\pi}{6}\rho d_2^3\, J \ ,$$

where, d_2 is the diameter of the 2 nm particle, and ρ is the density of the particle (taken as sulfuric acid) at ambient relative humidity.

Currently, despite the uncertainty associated with spatial variation, fine particle mass concentrations simulated from CMAQ are reasonably well matched to the measured concentrations (Arnold et al., 2003; Eder et al., 2003; Jun and Stein, 2004; Mebust et al., 2003; Park et al., 2005a; Tonnesen, 2003), but CMAQ's capability to reproduce aerosol number concentrations with reasonable accuracy is less well established. Some studies suggest that fine particle number and surface area concentrations are directly related to pulmonary disease (Donaldson et al., 1998; Penttinen et al., 2001; Woo et al., 2001b). These studies suggest that the particle number concentration itself is an important index in health study. Also, particle mass concentrations are not always correlated with particle number concentrations. For example, when a large number of ultra-fine particles are formed, the particle mass contributions do not significantly increase (Woo et al., 2001a). Hence, if particle number or surface area concentrations do, indeed, affect human health, strategies designed to meet NAAQS for $PM_{2.5}$ may be less effective for protecting human health (Woo et al., 2001a). Thus, a thorough evaluation of particle number concentrations should be performed, along with an analysis of ways to improve the accuracy of predicting simulated particle number concentrations. Here, three assumptions used by

CMAQ: assumed particle density, assumed size distribution of primary emissions, and assumed distribution of particle size, will be evaluated based on the measured data.

One of the applications of CMAQ includes source apportionment of $PM_{2.5.}$ Historically, source apportionment of $PM_{2.5}$ has been largely performed via several receptor-modeling methods that use a combination of statistical techniques and pollutant concentrations measured at a receptor. In particular, one of the most widely used receptor modeling techniques is the chemical mass balance (CMB) approach (Core et al., 1982; Friedlander, 1973; Watson et al., 2002b). In CMB, ambient chemical concentrations are expressed as the sum of products of source profiles and source contributions. This linear system of equations is solved for source contributions by weighted least square fitting. The CMB model makes the following six assumptions (US-EPA, 1990): 1. Compositions of source emissions are constant over the period of ambient and source sampling. 2. Chemical species do not react with each other (i.e., they add linearly). 3. All sources with a potential for contributing to the receptor have been identified and have had their emissions characterized. 4. The number of sources or source categories is less than or equal to the number of species. 5. Source profiles are linearly independent of each other. 6. Measurement uncertainties are random, uncorrelated, and normally distributed. Those six assumptions are never totally satisfied in actual practice (Watson et al., 2002b), and deviations from assumptions increase the uncertainty of the source contribution estimates. To overcome limitations in source apportionment using CMB, CMAQ is used to calculate mass contribution of each source. An analysis of the discrepancies of the mass contributions calculated by the source-oriented air quality and receptor models can be used to help improve the accuracy of results. A comparison will be done for two

complete months (July 2001 and January 2002) over eight stations in the Southeastern United States, and reasons for discrepancies are analyzed.

Sources of discrepancy in mass contributions calculated from CMAQ and those from CMB include inaccuracies in the simulation of $PM_{2.5}$ mass from CMAQ. The accuracy of simulated concentrations relies on the reliable inputs, and one of the most uncertain inputs to air quality models includes emissions. Four Dimensional Data Assimilation (FDDA) (Mendoza-Dominguez and Russell, 2000) has been developed to refine and evaluate emission inventories. The FDDA method links formal direct sensitivity analysis of three dimensional air quality models with an inverse modeling, which incorporates observational data of multiple species. FDDA can improve estimates of emission strengths of various source categories simultaneously and the processes of the model by adding information of the real state of the system. Fundamental assumptions of FDDA include that the major source for the discrepancies between the simulated and observed concentrations of pollutants is the emission inventory, and spatial allocation of emissions are less important to overall performance than the total mass of emissions (Mendoza-Dominguez and Russell, 2001b). The present study assesses the feasibility of using the FDDA approach to estimate regional emission strength adjustment over the continental United States for July 2001 and for January 2002. Regionally different emission scaling factors are estimated to capture the spatial variation of the systematic bias in emissions. Also, different scaling factors for weekdays and for weekends are calculated to analyze the day-of-week variation.

Structure and Scope of the Thesis

- **Chapter 2: Statistical analysis of $PM_{2.5}$ control strategy.** Required emission reductions for $PM_{2.5}$ in Atlanta, Georgia to meet the NAAQS are analyzed. The distributional property of $PM_{2.5}$ is used to account for the temporal variations of pollutant levels. The influence of correlations among $PM_{2.5}$ species on the estimated control levels is also investigated using the joint probability function.
- **Chapter 3: Uncertainty in air quality model evaluation for particulate matter due to spatial variation in pollutant concentrations.** Different spatial scales between simulated and measured concentrations causes uncertainty in model evaluation. Spatial variability of $PM_{2.5}$ individual species and total mass is analyzed. In addition, uncertainty in model evaluation due to the spatial variability of $PM_{2.5}$ is investigated using a data withholding method. Finally, a way to reduce the uncertainty in model performance evaluation is proposed.
- **Chapter 4: Evaluation of fine particle number concentrations in CMAQ.** Fine particle number concentrations simulated from CMAQ is investigated based on the measured particle number concentrations of 39 size bins from 3 nm to 3 μm for one year and eight months from 1/1/1999 to 8/31/2000 at an urban area located in Atlanta, Georgia. In addition, this chapter extensively analyzes the causes of errors, and examines the feasibility to improve the accuracy of the simulations.
- **Chapter 5: Comparison between chemical mass balance receptor and CMAQ model $PM_{2.5}$ source apportionment.** Mass contributions of $PM_{2.5}$ calculated using two fundamentally different air quality modeling approaches, source-based

and receptor-oriented, are compared. The source-based and receptor models selected are CMAQ and CMB, respectively. The analysis is done using $PM_{2.5}$ total mass and species concentrations over eight monitors in the Southeastern US in July 2001 and in January 2002. Sources of the discrepancy in source apportionment results from the two models are analyzed, and ways to improve the accuracy of the modeling results are investigated.

- **Chapter 6: Regional adjustment of emission strength via four-dimensional data assimilation.** The feasibility of using the FDDA approach to estimate regional emission strength adjustment over the continental United States is assessed. The emission sources adjusted are CO (total), NH3 (total), SO2 (total), NO_x (area/mobile/nonroad), NO_x (point), VOC (area/mobile/nonroad/point), VOC (biogenic), POA (total), PEC (total), and PMFINE (total). Regionally different emission scaling factors are estimated to capture the spatial variation of the systematic bias in emissions. In addition, different scaling factors for weekdays and for weekends are estimated to analyze the day-of-week variation of the emission bias.
- **Chapter 7: Conclusions and future research.** Elevated ambient particle levels affect human health, reduce visibility, deposit acid, and change Earth radiation balance. Atlanta, Georgia is one such area where $PM_{2.5}$ levels need to be controlled. This study estimated the controls required to meet the standard by analyzing the measured pollutant concentrations statistically. Another method to develop and assess control strategies is via three-dimensional photochemical air quality model. For the models to be used in such applications, a thorough

evaluation of these models is needed. In this thesis, the major sources of errors in current air quality models are analyzed, and ways to improve model performance are proposed.

References

Arnold, J. R., Dennis, R. L., and Tonnesen, G. S. (2003). Diagnostic evaluation of numerical air quality models with specialized ambient observations: testing the Community Multiscale Air Quality modeling system (CMAQ) at selected SOS 95 ground sites. *Atmospheric Environment* **37**, 1185-1198.

Blanchard, C. L., Carr, E. L., Collins, J. F., Smith, T. B., Lehrman, D. E., and Michaels, H. M. (1999). Spatial representativeness and scales of transport during the 1995 Integrated Monitoring Study in California's San Joaquin Valley. *Atmospheric Environment* **33**, 4775-4786.

Brauer, M., Ebelt, S. T., Fisher, T. V., Brumm, J., Petkau, A. J., and Vedal, S. (2001). Exposure of chronic obstructive pulmonary disease patients to particles: Respiratory and cardiovascular health effects. *Journal Of Exposure Analysis And Environmental Epidemiology* **11**, 490-500.

Butler, A. J., Andrew, M. S., and Russell, A. G. (2003). Daily sampling of PM2.5 in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. *Journal Of Geophysical Research-Atmospheres* **108**.

Byun, D. W., and Ching, J. K. (1999). Science algorithms of the EPA Models-3 Community Multiscale Air Qualtiy (CMAQ) modeling system. *Atmospheric modeling division, National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711*.

Carter, W. P. L. (1996). Condensed atmospheric photooxidation mechanisms for isoprene. *Atmospheric Environment* **24A**, 481-515.

Carter, W. P. L., Luo, D., and Malkina, I. L. (1997). Environmental Chamber Studies for Devalopment of an Updated Photochemical Mechanism for Reactivity Assessment. *Final Report for California Air Resources Board Contract No. 92-345, Coordniating Research Council, Inc., Project M-9 and National Renewable Energy Laboratory, Contract ZF-2-12252-07*.

Chan, C. C., and Hwang, J. S. (1996). Site representativeness of urban air monitoring stations. *Journal Of The Air & Waste Management Association* **46**, 755-760.

Chang, J. S., Brost, R. A., Isaksen, I. S. A., Madronich, S., Middleton, P., Stockwell, W. R., and Waleck, C. J. (1987). A three-dimensional Eulerian acid deposition model: physical concepts and formulation. *Journal Of Geophysical Research-Atmospheres* **92**, 14,681-14,700.

Core, J. E., Cooper, J. A., Hanrahan, P. L., and Cox, W. M. (1982). Particulate Dispersion Model Evaluation - A New Approach Using Receptor Models. *Journal Of The Air Pollution Control Association* **32**, 1142-1147.

Dennis, R. L. (1998). The environmental protection agency's third generation air quality modeling system: an overall perspective. *Proceedings of the American Meteorological Society 78th Annual Meeting, Phoenix, AZ, Jan. 11-16, 1998. 255-258*.

Dockery, D. W., and Pope, C. A. (1994). Acute Respiratory Effects of Particulate Air-Pollution. *Annual Review of Public Health* **15**, 107-132.

Dockery, D. W., Pope, C. A., Xu, X. P., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., and Speizer, F. E. (1993). An Association Between Air-Pollution And Mortality In 6 United-States Cities. *New England Journal Of Medicine* **329**, 1753-1759.

Donaldson, K., Li, X. Y., and MacNee, W. (1998). Ultrafine (nanometre) particle mediated lung injury. *Journal Of Aerosol Science* **29**, 553-560.

Donaldson, K., Stone, V., Seaton, A., and MacNee, W. (2001). Ambient particle inhalation and the cardiovascular system: Potential mechanisms. *Environmental Health Perspectives* **109**, 523-527.

Ebelt, S. T., Petkau, A. J., Vedal, S., Fisher, T. V., and Brauer, M. (2000). Exposure of chronic obstructive pulmonary disease patients to particulate matter: Relationships between personal and ambient air concentrations. *Journal Of The Air & Waste Management Association* **50**, 1081-1094.

Eder, B., Yu, S., and Dennis, R. (2003). An Evaluation of the Models-3 CMAQ Aerosol Module. *PM 2003 Meeting -- Particulate Matter: Atmospheric Sciences, Exposure and the Forth Colloquium on PM and Human Health, Pittsburgh, PA.*

Friedlander, S. (1973). Chemical Element Balances And Identification Of Air-Pollution Sources. *Environmental Science & Technology* **7**, 235-240.

Georgopoulos, P. G., and Seinfeld, J. H. (1982). Statistical Distributions Of Air Pollutant Concentrations. *Environmental Science & Technology* **16**, A401-&.

Gery, M. W., Whitten, G. Z., Killus, J. P., and Dodge, M. C. (1989). A photochemical kinetics mechanism for urban and regional scale computer modeling. *Journal Of Geophysical Research-Atmospheres* **94**, 12,925-12,956.

Gold, D. R., Damokosh, A. I., Pope, C. A., Dockery, D. W., McDonnell, W. F., Serrano, P., Retama, A., and Castillejos, M. (1999). Particulate and ozone pollutant effects on the respiratory function of children in Southwest Mexico City (vol 10, pg 8, 1999). *Epidemiology* **10**, 470-470.

Goldberg, M. S., Burnett, R. T., Valois, M. F., Flegel, K., Bailar, J. C., Brook, J., Vincent, R., and Radon, K. (2003). Associations between ambient air pollution and daily mortality among persons with congestive heart failure. *Environmental Research* **91**, 8-20.

Goswami, E., Larson, T., Lumley, T., and Liu, L. J. S. (2002). Spatial characteristics of fine particulate matter: Identifying representative monitoring locations in Seattle, Washington. *Journal Of The Air & Waste Management Association* **52**, 324-333.

Hadley, A., and Toumi, R. (2003). Assessing changes to the probability distribution of sulphur dioxide in the UK using a lognormal model. *Atmospheric Environment* **37**, 1461-1474.

Iwai, K., Mizuno, S., Miyasaka, Y., and Mori, T. (2005). Correlation between suspended particles in the environmental air and causes of disease among inhabitants: Cross-sectional studies using the vital statistics and air pollution data in Japan. *Environmental Research* **99**, 106-117.

Jun, M., and Stein, M. L. (2004). Statistical comparison of observed and CMAQ modeled daily sulfate levels. *Atmospheric Environment* **38**, 4427-4436.

Kruize, H., Hanninen, O., Breugelmans, O., Lebret, E., and Jantunen, M. (2003). Description and demonstration of the EXPOLIS simulation model: Two examples of modeling population exposure to particulate matter. *Journal Of Exposure Analysis And Environmental Epidemiology* **13**, 87-99.

Kulmala, M., Laaksonen, A., and Pirjola, L. (1998). Parameterizations for sulfuric acid/water nucleation rates. *Journal Of Geophysical Research-Atmospheres* **103**, 8301-8307.

Larsen, R. I. (1969). A new mathematical model of air pollutant concentration, averaging time and frequency. *Journal or the air pollution control association* **19**, 24.

Larson, T., Kalman, D., Wang, S., and Nothstein, G. (1989). Urban air toxics mitigation study, phase 1. *Technical report, Puget Sound Air Pollution Control Agency, Seattle, Washington, 1989.*

Latha, K. M., and Badarinath, K. V. S. (2005). Shortwave radiative forcing efficiency of urban aerosols - a case study using ground based measurements. *Chemosphere* **58**, 217-220.

Lu, H. C. (2002). The statistical characters of PM10 concentration in Taiwan area. *Atmospheric Environment* **36**, 491-502.

Lu, H. C., and Fang, G. C. (2002). Estimating the frequency distributions of PM10 and PM2.5 by the statistics of wind speed at Sha-Lu, Taiwan. *Science Of The Total Environment* **298**, 119-130.

McNair, L. A., Harley, R. A., and Russell, A. G. (1996). Spatial inhomogeneity in pollutant concentrations, and their implications for air quality model evaluation. *Atmospheric Environment* **30**, 4291-4301.

Mebust, M. R., Eder, B. K., Binkowski, F. S., and Roselle, S. J. (2003). Models-3 community multiscale air quality (CMAQ) model aerosol component - 2. Model evaluation. *Journal Of Geophysical Research-Atmospheres* **108**.

Mendoza-Dominguez, A., and Russell, A. G. (2000). Iterative inverse modeling and direct sensitivity analysis of a photochemical air duality model. *Environmental Science & Technology* **34**, 4974-4981.

Mendoza-Dominguez, A., and Russell, A. G. (2001). Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. *Atmospheric Environment* **35**, 2879-2894.

Park, S.-K., Cobb, C. E., Wade, K., Mulholland, J., Hu, Y., and Russell, A. (2005). Uncertainty in Air Qualiy Model Evaluation for Particulate Matter due to Spatial Variation in Pollutant Concentrations. *Atmospheric Environment, submitted for publication*.

Penttinen, P., Timonen, K. L., Tittanen, P., Mirme, A., Ruuskanen, J., and Pekkanen, J. (2001). Number concentration and size of particles in urban air: Effects on Spirometric lung function in adult asthmatic subjects. *Environmental Health Perspectives* **109**, 319-323.

Pope, C. A., Thun, M. J., Namboodiri, M. M., Dockery, D. W., Evans, J. S., Speizer, F. E., and Heath, C. W. (1995). Particulate Air-Pollution As A Predictor Of Mortality In A Prospective-Study Of Us Adults. *American Journal Of Respiratory And Critical Care Medicine* **151**, 669-674.

Pope, C. A., Verrier, R. L., Lovett, E. G., Larson, A. C., Raizenne, M. E., Kanner, R. E., Schwartz, J., Villegas, M., Gold, D. R., and Dockery, D. W. (1999). Heart rate variability associated with particulate air pollution. *American Heart Journal* **138**, 890-899.

PSU/NCAR (2003). PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3. *Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research.*

Restrepo, C., Zimmerman, R., Thurston, G., Clemente, J., Gorczynski, J., Zhong, M. H., Blaustein, M., and Chen, L. C. (2004). A comparison of ground-level air quality data with New York State Department of Environmental Conservation monitoring

stations data in South Bronx, New York. *Atmospheric Environment* **38**, 5295-5304.

Russell, A., and Dennis, R. (2000). NARSTO critical review of photochemical models and modeling. *Atmospheric Environment* **34**, 2283-2324.

Saldiva, P. H. N., Pope, C. A., Schwartz, J., Dockery, D. W., Lichtenfels, A. J., Salge, J. M., Barone, I., and Bohm, G. M. (1995). Air-Pollution And Mortality In Elderly People - A Time-Series Study In Sao-Paulo, Brazil. *Archives Of Environmental Health* **50**, 159-163.

Satheesh, S. K., and Moorthy, K. K. (2005). Radiative effects of natural aerosols: A review. *Atmospheric Environment* **39**, 2089-2110.

Sheppard, L., Levy, D., and Checkoway, H. (2001). Correcting for the effects of location and atmospheric conditions on air pollution exposures in a case-crossover study. *Journal Of Exposure Analysis And Environmental Epidemiology* **11**, 86-96.

Stockwell, W. R., Middleton, P., Chang, J. S., and Tang, X. (1990). The second generation regional acid deposition model chemical mechanism for regional air quality modeling. *Journal Of Geophysical Research-Atmospheres* **95**, 16,343-16,367.

Takemura, T., Nozawa, T., Emori, S., Nakajima, T. Y., and Nakajima, T. (2005). Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model. *Journal Of Geophysical Research-Atmospheres* **110**.

Tonnesen, G. S. (2003). Regional haze modeling: Recent modeling results for VISTAS and WRAP. *Models-3 User's workshop, Research Triangle Park, North Carolina*.

Ulrich, M. M. W., Alink, G. M., Kumarathasan, P., Vincent, R., Boere, A. J. F., and Cassee, F. R. (2002). Health effects and time course of particulate matter on the cardiopulmonary system in rats with lung inflammation. *Journal Of Toxicology And Environmental Health-Part A* **65**, 1571-1595.

US-EPA (1990). CMB8 User's Mannual. *Office of Air Quality, Planning and Standards, Research Triangle Park, NC 27711*.

US-EPA (2004a). Air Quality Criteria for Particulate Matter (October 2004). *EPA 600/P-99/002aF-bF*.

US-EPA (2004b). SMOKE User's manual. *http://cf.unc.edu/cep/empd/products/smoke*.

Vedal, S. (1997). Ambient particles and health: Lines that divide. *Journal Of The Air & Waste Management Association* **47**, 551-581.

Watson, J. G., Zhu, T., Chow, J. C., Engelbrecht, J., Fujita, E. M., and Wilson, W. E. (2002). Receptor modeling application framework for particle source apportionment. *Chemosphere* **49**, 1093-1136.

Wilemski, G. (1984). Composition Of The Critical Nucleus In Multicomponent Vapor Nucleation. *Journal Of Chemical Physics* **80**, 1370-1372.

Woo, K. S., Chen, D. R., Pui, D. Y. H., and McMurry, P. H. (2001a). Measurement of Atlanta aerosol size distributions: Observations of ultrafine particle events. *Aerosol Science And Technology* **34**, 75-87.

Woo, K. S., Chen, D. R., Pui, D. Y. H., and Wilson, W. E. (2001b). Use of continuous measurements of integral aerosol parameters to estimate particle surface area. *Aerosol Science And Technology* **34**, 57-65.

CHAPTER 2

STATISTICAL ANALYSIS OF $PM_{2.5}$ CONTROL LEVELS IN METROPOLITAN ATLANTA, GEORGIA

(S.-K. Park, K. Manomaiphiboon, A.G. Russell.

Atmospheric Environment, in preparation)

Abstract

Evidences continue to mount linking elevated $PM_{2.5}$ levels with adverse health effects. In response, the National Ambient Air Quality Standard (NAAQS) for $PM_{2.5}$ was promulgated in 1997 (US-EPA, 2004a). In metropolitan Atlanta, Georgia, $PM_{2.5}$ levels exceed the annual NAAQS (15 μg m^{-3}). The goal of this study has been to calculate emission reductions required to reduce $PM_{2.5}$ levels in Atlanta below the NAAQS, and quantify how specific sources impact $PM_{2.5}$ levels. Emission reductions were calculated based on the rollback method using $PM_{2.5}$ mass and species concentrations measured at four monitors located in Atlanta: Fort McPherson (FTM), South DeKalb (SDK), Tucker (TUC), and Jefferson Street (JST) from March 31, 1999, to December 31, 2003. A distributional property of $PM_{2.5}$ is used to calculate the level of uncertainty meeting the NAAQS with a given emission reductions. Results found that the required emission reductions to meet the standard were 22% (FTM), 18% (SDK), 20% (TUC), and 20% (JST) at 50% confidence interval (CI). The reductions were 30% (FTM), 21% (SDK), 42% (TUC), and 32% (JST) at 95% CI. Reduction in the emissions for each $PM_{2.5}$ species is also calculated using the Koehler-Symanowski (Koeh-Syma) joint pdf to

account for the correlation between each species. The analysis is done for two stations, SDK and JST, because $PM_{2.5}$ species data from other stations did not fit KS-pdf. Reductions for sulfate and organic carbon meeting the NAAQS decreases up to 8% for SDK and 20% for JST. The results in these analyses can be used as a guideline in establishing the $PM_{2.5}$ control strategy in the Atlanta area.

Key words: $PM_{2.5}$; NAAQS; Emissions; Probability Density Function; Control Strategy

2.1 Introduction

PM is frequently the most obvious form of air pollution because it reduces visibility and soils surfaces (Wark et al., 1997). In 1997, the National Ambient Air Quality Standard (NAAQS) for $PM_{2.5}$ (particulate matter with an aerodynamic diameter of less than 2.5 μm) were promulgated in response to an increasing number of scientific studies linking elevated fine particle concentrations with adverse health effects. The long-term standard set the allowable three-year average of the annual mean $PM_{2.5}$ concentrations at less than or equal to 15 $\mu g/m^3$; the short-term standard sets the three-year average of the 98^{th} percentile of the 24-hour $PM_{2.5}$ concentrations at 65 $\mu g/m^3$. Because of the adverse health effects of $PM_{2.5}$, its control strategies are important and need to be addressed. Metropolitan Atlanta, Georgia is one such area in which the annual $PM_{2.5}$ level is higher than the annual NAAQS, so the control of $PM_{2.5}$ is necessary. The goal of this paper is to quantify the required emission reductions for $PM_{2.5}$ to meet the annual NAAQS in Atlanta. The property that $PM_{2.5}$ data follow a specified distribution is used in developing the control strategy.

The distributional property of pollutant concentrations has been studied for decades. Multiple statistical distributions of air pollutant concentrations have been reviewed (Georgopoulos and Seinfeld, 1982), and a computer program has been developed for fitting statistical distributions to air pollution data using maximum likelihood estimation (Holland and Fitzsimons, 1982). The accuracy of the statistical distribution fit to the extreme vales of pollutant concentrations was also analyzed, and the result revealed that the very extreme values have large uncertainties and are very sensitive to the choice of the distribution (Chock and Sluchak, 1986). These distributions studied are widely used in solving environmental problems.

The frequency distributions of the source contributions of PM_{10} in the South Coast Air Basin (SoCAB) were used to analyze the different sources that have similar chemical profiles (Kao and Friedlander, 1995). The statistical distribution of personal exposure to PM_{10} levels was modeled to quantify how much the ambient concentrations affect the human exposure (Ott et al., 2000). Results showed that when the ambient concentrations were controlled, the median of the personal exposure to PM_{10} ranged from 32.0 $\mu g/m^3$ (Toronto) to 34.4 $\mu g/m^3$ (Phillipsburg) to 48.8 $\mu g/m^3$ (Riverside). The probability density function (pdf) was used for PM_{10} concentrations measured in Taiwan to estimate the probabilities of exceeding the air quality standard and to determine the emission source reductions of PM_{10} concentrations necessary to meet the air quality standard (Lu, 2002). Microenvironment concentrations and the contributions of indoor sources were approximated by a lognormal distribution, and the time spent in a microenvironment and the penetration factor were approximated by a beta distribution (Kruize et al., 2003). The pdf of air pollutants (e.g., NO_2, NO_x, benzene, and O_3) was also

used to estimate bi-variate distributions of concentrations, concentration roses, and the probability that when concentrations of one pollutant reach an extreme, another will be present in an extreme concentration as well (Genikhovich et al., 2005). Because of the usefulness of the frequency distribution of pollutant concentrations, a statistical method to predict the frequency distribution of PM_{10} and $PM_{2.5}$ at specific wind speeds was developed with measured data collected at the Sha-Lu station in Taiwan (Lu and Fang, 2002). The temporal change over 40 years in the probability distribution of sulfur dioxide at 10 monitors in the United Kingdom was also analyzed (Hadley and Toumi, 2003). Moreover, the possibilities of developing a general probability model for fitting environmental quality data have been explored (Singh et al., 2001). Errors associated with the fitted distribution, when less frequently sampled data were used, also have been analyzed (Rumburg et al., 2001) .

Here, the property that pollutant concentrations follow specific underlying distributions is used to account for the temporal variations of pollutant levels when control strategies are developed. In addition, the Koehler-Symanowski (Koehler and Symanowski, 1995) joint pdf is used to estimate the required emissions. The benefit of using joint pdf is that the correlation between $PM_{2.5}$ species is taken into account for estimating the needed emission reductions. The analysis is performed using daily average $PM_{2.5}$ mass and species concentrations measured from 1999 to 2003 in Atlanta. Monitoring data are used from four locations: Fort McPherson (FTM), South DeKalb (SDK), Tucker (TUC), and Jefferson Street (JST). These data are collected as a part of the Assessment of Spatial Aerosol Composition in Atlanta (ASACA) project and the

SouthEastern Aerosol Research and Characterization (SEARCH) study (Butler, 2000; Butler et al., 2003; Hansen et al., 2003) (Table 2-1 and Figure 2-1).

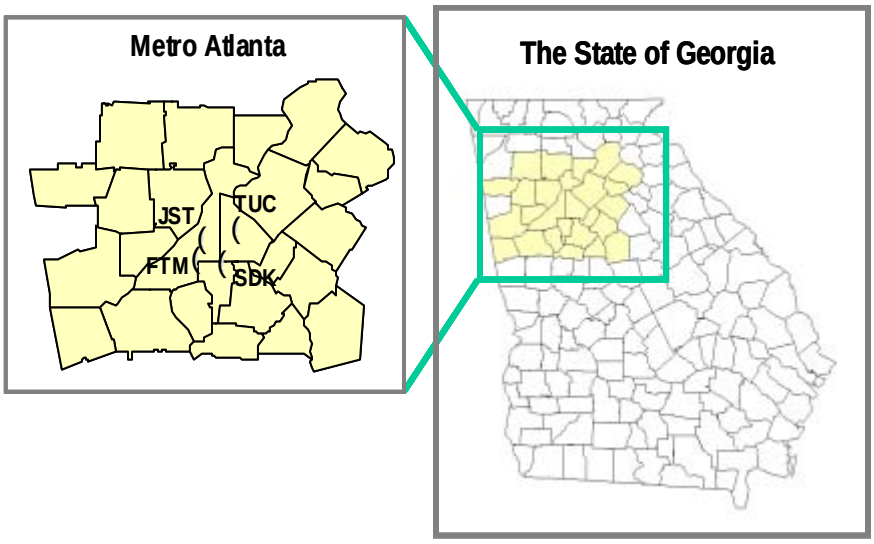


Figure 2-1. ASACA $PM_{2.5}$ monitoring stations.

Table 2-1. Site characteristics (Butler, 2000; Butler et al., 2003)

Station	Background	Latitude, Longitude
Jefferson Street	Urban, industrial	33.78 N, 84.41 W
Fort McPherson	Urban, near major highway	33.70 N, 84.44 W
South DeKalb	Residential, near major highway	33.70 N, 84.29 W
Tucker	Suburban/commercial	33.85 N, 84.21 W

2. Methods

2.2.1. Univariate probability density functions

The univariate probability density functions (pdfs) used in this study are selected from among three distributions, lognormal, Weibull, and gamma distributions, which have been proven particularly useful in representing air quality data (Georgopoulos and Seinfeld, 1982). Probability density and cumulative distribution functions (cdf) for lognormal, Weibull, and gamma distributions are summarized in Table 2-2. Parameters of the distributions were estimated using the method of the moments (MM) and the maximum likelihood estimation (MLE) (Table 2-3).

Table 2-2. Probability density and cumulative distribution functions, and means and variances for the lognormal, Weibull, and gamma distributions

Distribution	pdf	cdf	mean	variance
Lognormal	$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma\,x}e^{\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right)}$	$F(x)=\Phi\left(\frac{\ln x-\mu}{\sigma}\right)$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$
Weibull	$f(x)=\frac{a}{b^a}x^{a-1}e^{\left(-\left(\frac{x}{b}\right)^a\right)}$	$F(x)=1-e^{\left(-\left(\frac{x}{b}\right)^a\right)}$	$b\Gamma(1+\frac{1}{a})$	$b^2\left[\Gamma(1+\frac{2}{a})-\Gamma^2(1+\frac{1}{a})\right]$
Gamma	$f(x)=\frac{x^{a-1}}{b^a\Gamma(a)}e^{\left(-\left(\frac{x}{b}\right)\right)}$	$F(x)=\gamma(\frac{x}{b},a)$	ab	ab^2

Table 2-3. Equations for estimating parameters of the distributions in Table 2-2.

Distribution	MM	MLE
Lognormal	$\hat{\mu} = 2\ln M_1 - \frac{1}{2}\ln M_2 \quad \hat{\sigma}^2 = \ln M_2 - 2\ln M_1$	$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln x_i \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(\ln x_i - \hat{\mu})^2$
Weibull	$\frac{\Gamma(1+\frac{2}{\hat{a}})}{\Gamma^2(1+\frac{1}{\hat{a}})} = \frac{M_2}{M_1^2} \quad \hat{b} \quad \frac{M_1}{\Gamma(1+\frac{1}{\hat{a}})}$	$\hat{a} = \left[\left(\frac{\sum_{i=1}^{n} x_i^{\hat{a}}\ln x_i}{\sum_{i=1}^{n} x_i^{\hat{a}}}\right) - \frac{1}{n}\sum_{i=1}^{n}\ln x_i\right]^{-1} \quad \hat{b} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^{\hat{a}}\right)^{\frac{1}{\hat{a}}}$
Gamma	$\hat{b} = \frac{M_2 - M_1^2}{M_1} \quad \hat{a} = \frac{M_1}{\hat{b}}$	$\frac{\Gamma'(\hat{a})}{\Gamma(\hat{a})} = \ln\left(\frac{\hat{a}}{n}\sum_{i=1}^{n} x_i\right) - \frac{1}{n}\sum_{i=1}^{n}(\ln x_i) \quad \hat{b} = \frac{1}{n\hat{a}}\sum_{i=1}^{n} x_i$

The MM estimates the parameters of a distribution using moments. The general form of the m[th] moments for the distribution is

$$M_m = \int_0^{\infty} (x_i)^m \, p(x_i)\, dx_i \tag{2-1}$$

where $p(x_i)$ is the pdf of the distribution. When a discrete data set of size *n* is considered, the m^{th} moment is

$$M_m = \frac{1}{n}\sum_{i=1}^{n} x_i^{\,m} \tag{2-2}$$

The mean of sample data, *E(x)*, corresponds to the first moment (M_1),

$$E(x) = M_1 \tag{2-3}$$

and the variance of the sample data, *Var(x)*, is

$$Var(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - M_1)^2 \approx \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2x_i M_1 + M_1^2) = M_2 - M_1^2 \tag{2-4}$$

Parameters of distributions are obtained using the expectation and variance of the distribution (Table 2-2). Parameters of the lognormal, Weibull, and gamma distributions, estimated from MM, are summarized in Table 2-3.

The MLE estimates the parameters of a distribution that maximize the likelihood function, defined as the joint pdf of the observations in a random sample size of *n*

$$L(\theta_1, \ldots, \theta_k) = \prod_{i=1}^{n} p(x_i ; \theta_1, \ldots, \theta_k), \tag{2-5}$$

where $\theta_1, \ldots, \theta_k$ are the parameters of the distribution. Parameters maximizing the likelihood function can be calculated by partial derivatives of L with respect to each parameter, setting the derivatives equal to zero.

The goodness-of-fit, the index estimating how well the distribution fits the raw data, is checked using two statistical tests: chi-square (χ^2) and Kolmogorov-Smirnov (Kol-Smir). The chi-square test compares the data histogram with the pdf, so the test statistic involves comparing counts of data values falling into each class in relation to the computed theoretical probabilities,

$$\chi^2 = \sum_{classes} \frac{(\#observed - \#expected)^2}{\#expected} = \sum_{classes} \frac{(\#observed - n \Pr\{data\ in\ class\})^2}{n \Pr\{data\ in\ class\}} \tag{2-11}$$

The chi-square test statistic is dependent on the number of the class, and it is recommended to have at least five samples in each class. This analysis divided data into seven classes. Smaller chi-square statistics show better agreement between the empirical and theoretical distributions. The critical point of the chi-square test statistic can be

calculated using the cumulative chi-square distribution when the degrees of freedom (= # of classes – # of parameters fit –1) and the confidence interval are given.

The Kol-Smir test consists of comparing the empirical and theoretical cdfs. The Kol-Smir test is

$$D_n = max \mid F_n(x) - F(x) \mid \qquad (2\text{-}12)$$

where $F_n(x)$ is the empirical cumulative probability, estimated as $F_n(x_i) = \frac{i}{n}$ for the i^{th} smallest data value, and $F(x)$ is the theoretical cdf evaluated at x. Smaller Kol-Smir test statistics show better agreement between the empirical and theoretical distributions. The critical point of the Kol-Smir test statistics for the 95% confidence interval (CI) is approximated by $\frac{1.38}{\sqrt{n}}$, where n is the number of samples.

2.2.2. Bivariate probability density function

The bivariate pdf used in this study is the Koehler Symanowski (Koeh-Syma) pdf (Koehler and Symanowski, 1995). Bivariate Koeh-Syma pdf has been applied to the atmospheric wind fluctuations as well (Manomaiphiboon and Russell, 2003). Here, bivariate KS-pdfs are applied to $PM_{2.5}$ species concentrations (sulfate and ammonium; organic carbon, OC and elemental carbon, EC). The advantage of using KS-pdfs for air quality data is that it incorporates correlations between datasets, and strictly conserves the original shape of each marginal distribution. Thus, KS-pdfs are one of the more useful joint pdfs for $PM_{2.5}$ species, which are correlated with each other and follow underlying univariate distributions.

Let $X \equiv (X_1, X_2, \ldots, X_n)$ be a multivariate of n components and $x \equiv (x_1, x_2, \ldots, x_n)$ be the real-valued vector of X. Let $F \equiv F(\boldsymbol{x}) \equiv F_{\boldsymbol{x}}(\boldsymbol{x})$ be the cdf of X and $F_{X_i} \equiv F_{X_i}(x_i)$ be the marginal cdf of univariate X_i. Also, let $f \equiv f(\boldsymbol{x}) \equiv f_{\boldsymbol{x}}(\boldsymbol{x})$ be the pdf of X and $f_{X_i} \equiv f_{X_i}(x_i)$ be the marginal pdf of X_i. The KS-pdf is given by (Koehler and Symanowski, 1995):

$$f = f_{X_1} f_{X_2} D_1 D_2 C_{12}^{-\alpha_{12}} \left(1 + \alpha_{12} \alpha_{1+}^{-1} \alpha_{2+}^{-1} D_1^{-1} D_2^{-1} C_{12}^{-2} F_{X_1}^{1/\alpha_{1+}} F_{X_2}^{1/\alpha_{2+}}\right), \quad (2\text{-}13)$$

where $C_{12} = F_{X_1}^{1/\alpha_{1+}} + F_{X_2}^{1/\alpha_{2+}} - F_{X_1}^{1/\alpha_{1+}} F_{X_2}^{1/\alpha_{2+}}$,

$D_1 = \alpha_{1+}^{-1}[\,\alpha_{11} + (\alpha_{12} F_{X_2}^{1/\alpha_{2+}} C_{12}^{-1})\,]$,

$D_2 = \alpha_{2+}^{-1}[\,\alpha_{22} + (\alpha_{12} F_{X_1}^{1/\alpha_{2+}} C_{12}^{-1})\,]$,

$\alpha_{1+} = \alpha_{11} + \alpha_{12}$, and

$\alpha_{2+} = \alpha_{12} + \alpha_{22}$.

In this study, parameters of the distribution, α_{ij}, are estimated by the MLE using the sequential quadratic programming (SQP) method (Fletcher, 2003). The goodness-of-fit is checked using the boostrap method with a 95% CI.

2.3 Results and Discussion

Speciated $PM_{2.5}$ concentrations in Atlanta from March 1, 1999 to December 31, 2003 are summarized in Figure 2-2. Average measured $PM_{2.5}$ levels in Atlanta for that period are 19.0 μg m^{-3} (FTM), 18.2 μg m^{-3} (SDK), 18.7 μg m^{-3} (TUC), and 18.6 μg m^{-3}

(JST). Sulfate and OC are the two major $PM_{2.5}$ species, which account for 50% of $PM_{2.5}$ mass. Spatial variability was analyzed using the ANalysis Of VAriance (ANOVA) (Figure 2-3). Mean $PM_{2.5}$ concentrations in Figure 2-3 are slightly different from those in Figure 2-2 because spatial variability was analyzed only on days when measured concentrations are available for all four stations. The "p" value in Figure 2-3 is compared against α, which is 0.05 for a confidence level of 95 %. If "p" is larger than α, the null hypothesis that pollutant levels from four different sites are equal is accepted. Thus, from a statistical point of view, there was no significant spatial variation in $PM_{2.5}$ mass in Atlanta. The low variation implies that the pollutant concentrations from these monitors are spatially representative (Park et al., Submitted). Temporal variations of $PM_{2.5}$ were also analyzed (Figures 2-4 and 2-5). Apparent seasonal variation is observed for sulfate, nitrate, and ammonium concentrations. Pollutant concentrations on weekdays are slightly higher than those in weekends. This trend is consistent with what has been used in the temporal profiles of emissions (US-EPA, 2004e).

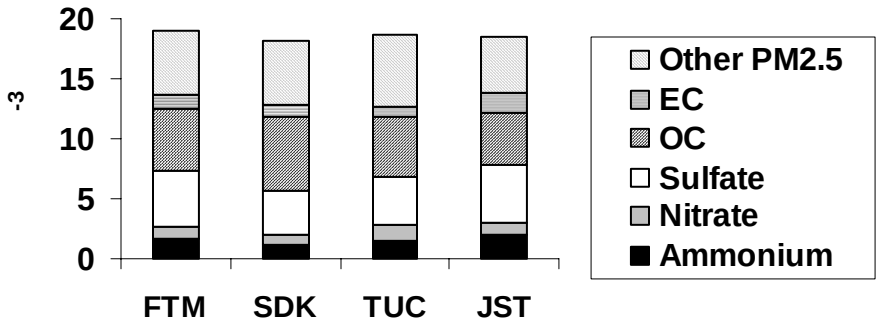


Figure 2-2. Average $PM_{2.5}$ species concentrations from 1999 to 2003.

Analysis of Variance

Source	DF	SS	MS	F	P
Factor	3	540.1	180.0	2.61	0.050
Error	1188	81965.3	69.0		
Total	1191	82505.4			

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
FTM	298	19.444	8.175	(------*------)
SDK	298	19.574	8.394	(------*------)
TUC	298	20.241	9.201	(------*------)
JST	298	18.365	7.352	(------*------)
Pooled StDev =		8.306		18.0 19.2 20.4 21.6

Figure 2-3. ANalysis Of VAriance (ANOVA) for $PM_{2.5}$ in Atlanta from 1999 to 2003.

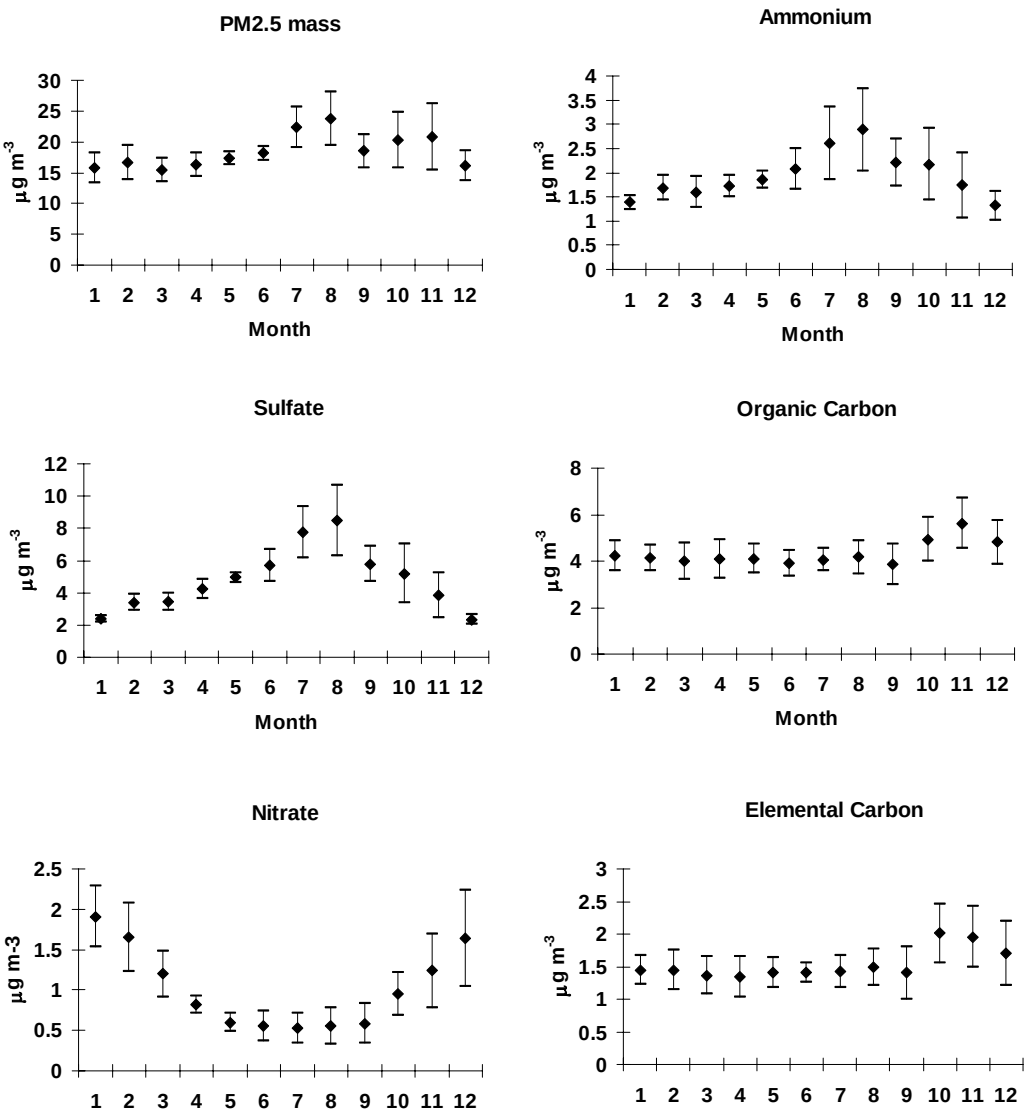


Figure 2-4. Monthly average $PM_{2.5}$ concentrations and standard deviations at Jefferson Street (JST) station in Atlanta from 1999 to 2003.

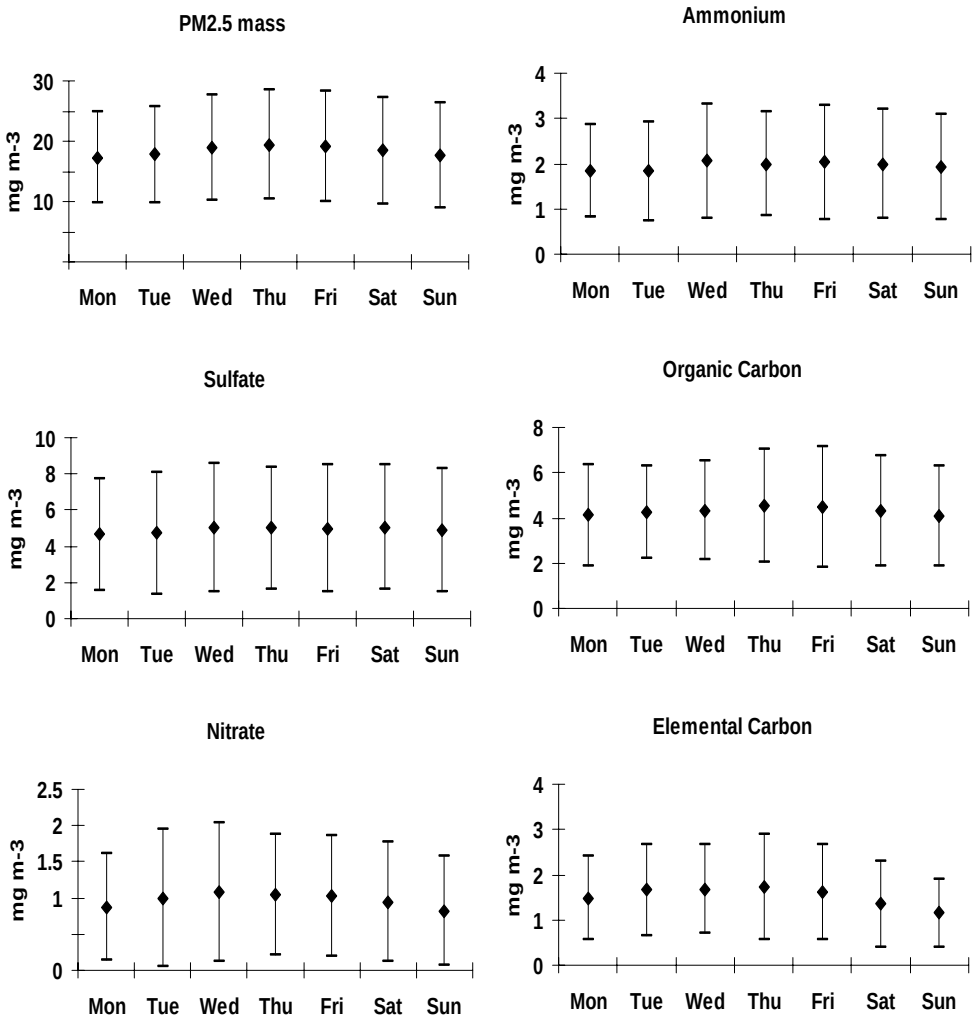


Figure 2-5. Day-of-week average $PM_{2.5}$ concentrations and standard deviation at JST from 1999 to 2003.

$PM_{2.5}$ levels in Atlanta exceed the long-term standard of 15 μg m^{-3}, but not the 24-hour standard of 65 μg m^{-3} (Figure 2-6). When annual mean concentrations are a

concern, the required emission reduction are often calculated using the rollback equation (de Nevers et al., 1977; Georgopoulos and Seinfeld, 1982).

$$R = \frac{(E(c) - c_b) - (E(c)_s - c_b)}{E(c) - c_b} = \frac{E(c) - E(c)_s}{E(c) - c_b}, \quad (2\text{-}14)$$

where, $E(c)$ is the annual average pollutant concentrations, and c_b is the background concentration (~ 0.8 μg m^{-3}) (Baumann et al., 2005). $E(c)_s$ is the air quality standard for annual $PM_{2.5}$ mass (15 μg m^{-3}), and R is the emission reduction required to meet the standard.

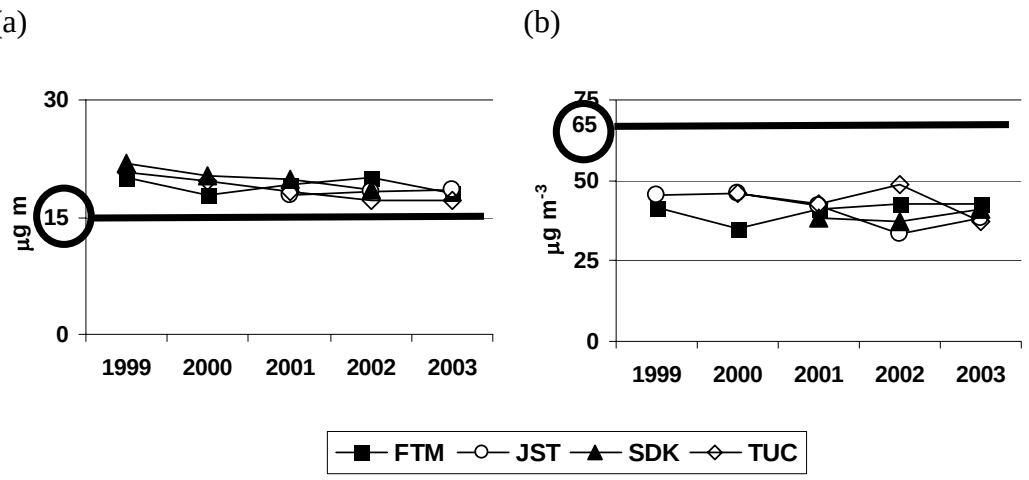
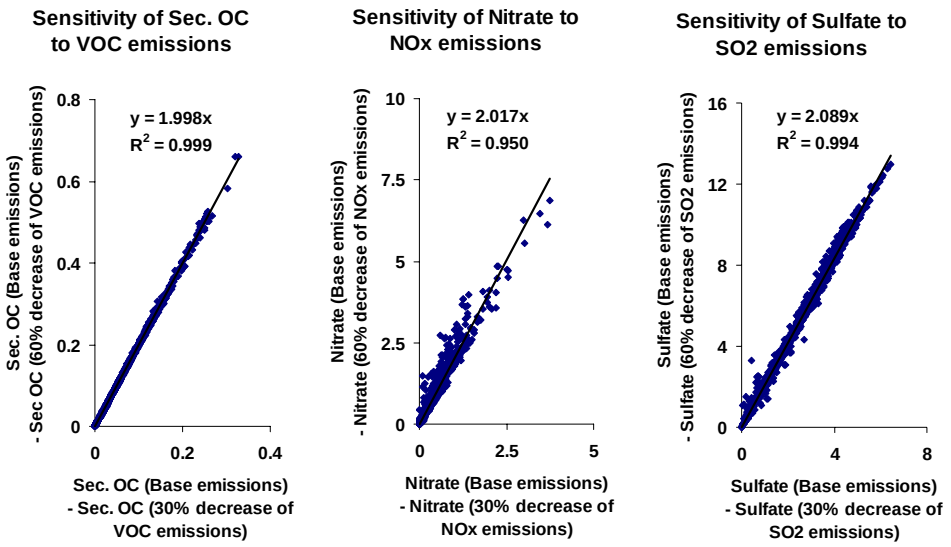


Figure 2-6. (a) Annual average $PM_{2.5}$ mass concentrations (b) 98th percentile of daily average $PM_{2.5}$ mass from 1999 to 2003.

The rollback method is valid when pollutant concentrations respond linearly to the overall emissions. However, secondary pollutant concentrations may not respond linearly to emission strengths. Responses of major secondary pollutants concentrations (e.g., sulfate, nitrate, and secondary organic carbon) to emission strengths (e.g., SO2, NOx, and VOC emissions) were checked using air quality modeling for July 2001 and January 2002 in a separate study (Park and Russell, 2003b). Community Multiscale Air Quality (CMAQ) model was run three times; using base level emissions, after 30% of emissions are reduced, and after 60% of emissions are reduced. The difference between pollutant levels with base emissions and those with 60% reduced emissions are twice of the difference between pollutant levels with base emissions and those with 30% reduced emissions (Figure 2-7). This result suggest that major secondary species of $PM_{2.5}$ can be assumed to respond linearly to precursor emissions

(a)



(b)

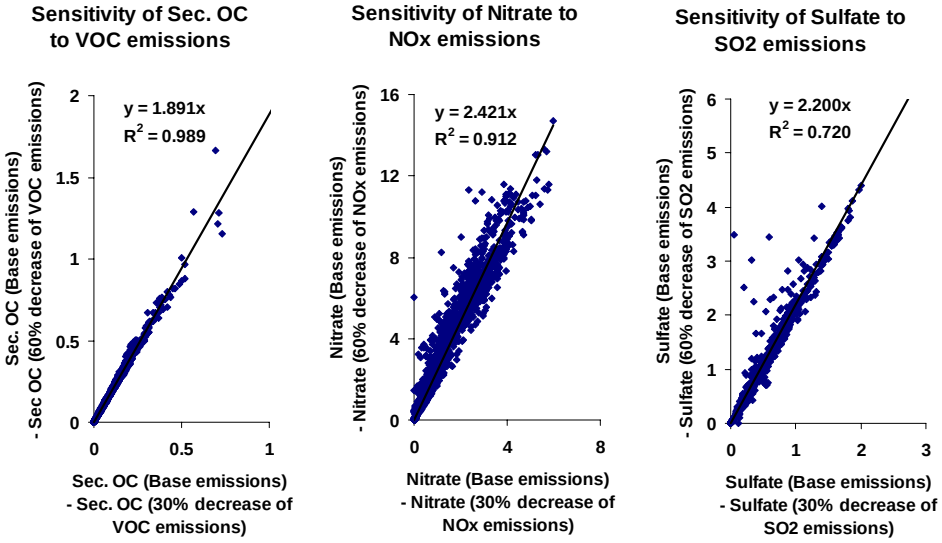


Figure 2-7. Sensitivity of secondary species of $PM_{2.5}$ to emissions in (a) July 2001 and (b) January 2002.

The amount of the emission reduction required to meet the annual standard is 22% (FTM), 18% (SDK), 21% (TUC), and 20% (JST) based on equation (2-14) and the average of the annual mean $PM_{2.5}$ mass for five years (Figure 2-2). The amount of the reduction is calculated assuming that the future emission strength is the same as the present emission strength in the absence of control. Thus, the calculation was done based on the average of the annual mean concentrations. However, future emission strengths vary, so the effectiveness of the $PM_{2.5}$ control is uncertain if emission reductions are calculated based on the average of the annual mean concentrations. The variation of annual mean concentrations cannot be calculated via annual mean concentrations alone because the number of data is not sufficient to estimate the parameters of the distribution. Thus, the abundant daily $PM_{2.5}$ data are used to estimate the temporal variation of the annual mean $PM_{2.5}$ level using the distributional property of $PM_{2.5}$ concentrations. The underlying distributional fit to the daily $PM_{2.5}$ level is selected, and the average of 365 randomly selected daily $PM_{2.5}$ concentrations is used to derive an annual mean concentration.

Raw daily $PM_{2.5}$ data are fitted among lognormal, Weibull, and gamma distributions separately for each station and for each year using the Maximum Likelihood Estimation (MLE), and the Method of Moments (MM). The best-fitted distribution is selected based on the goodness of fit test (Table 2-4).

Table 2-4. Estimated parameters for lognormal, Weibull, and gamma distributions of daily $PM_{2.5}$ mass concentrations from 1999 to 2003.

			JST 1999 276	FTM 1999 211	TUC 1999 272	JST 2000 298	FTM 2000 191	TUC 2000 224	JST 2001 279	FTM 2001 194	SDK 2001 252
Lognormal	MM	μ	2.96	2.90	2.99	2.97	2.80	2.91	2.96	2.86	2.79
		σ	0.46	0.46	0.44	0.43	0.39	0.45	0.44	0.42	0.43
	MLE	μ	2.95	2.88	2.98	2.97	2.80	2.91	2.95	2.85	2.78
		σ	0.50	0.50	0.47	0.43	0.40	0.45	0.46	0.47	0.44
Weibull	MM	a	2.2	2.2	2.2	2.4	2.7	2.3	2.3	2.4	2.4
		b	24.3	22.7	24.6	24.3	19.9	22.9	23.8	21.5	20.1
	MLE	a	2.2	2.2	2.3	2.4	2.7	2.3	2.3	2.4	2.4
		b	24.3	22.8	24.8	24.3	20.0	22.9	23.9	21.6	20.1
Gamma	MM	a	4.3	4.2	4.6	5.0	6.2	4.5	4.8	5.1	4.9
		b	5.0	4.7	4.7	4.3	2.8	4.5	4.4	3.8	3.6
	MLE	a	4.3	4.3	4.8	5.5	6.6	5.2	5.0	5.2	5.3
		b	5.0	4.7	4.6	3.9	2.7	3.9	4.2	3.7	3.3

			TUC 2001 237	JST 2002 342	FTM 2002 224	SDK 2002 362	TUC 2002 353	JST 2003 344	FTM 2003 345	SDK 2003 361
Lognormal	MM	μ	2.84	2.76	2.92	2.82	2.84	2.75	2.80	2.83
		σ	0.54	0.40	0.41	0.43	0.39	0.44	0.45	0.42
	MLE	μ	2.81	2.75	2.91	2.82	2.84	2.73	2.79	2.82
		σ	0.61	0.43	0.42	0.41	0.41	0.48	0.46	0.45
Weibull	MM	a	1.8	2.5	2.5	2.4	2.6	2.3	2.2	2.4
		b	22.3	19.2	22.7	20.8	20.8	19.4	20.4	20.9
	MLE	a	1.8	2.6	2.4	2.3	2.6	2.3	2.2	2.4
		b	22.4	19.2	22.7	20.7	20.9	19.4	20.6	20.9
Gamma	MM	a	3.0	5.8	5.3	4.8	5.9	4.6	4.5	5.2
		b	6.7	3.0	3.8	3.8	3.1	3.7	4.1	3.5
	MLE	a	3.0	5.8	5.9	6.0	6.3	4.8	4.9	5.4
		b	6.5	2.9	3.4	3.0	3.0	3.6	3.7	3.4

Data for SDK in 1999 and 2000 and data for TUC in 2003 were not available. Thus, 17 datasets were fit to the distributions. Goodness-of-fit was checked using both

Kolmogorov-Smirnov (K-S) and chi-square test statistics (Tables 2-5 and 2-6). The K-S test statistic (D_n) is expressed as

$$D_n = max \mid F_n(x) - F(x) \mid$$

where $F_n(x)$ is the empirical cumulative probability, estimated as $F_n(x_i) = \frac{i}{n}$ for the i^{th} smallest data value; and F(x) is the theoretical cumulative distribution function (CDF) evaluated at x. Thus, the K-S test statistic (D_n) looks for the largest difference, in an absolute value sense, between the empirical and fitted CDFs. The Chi square (χ^2) test statistic involves the counts of data values falling into each class in relation to the computed theoretical probabilities,

$$\chi^2 = \sum_{classes} \frac{(\#observed - \#expected)^2}{\#expected}$$

$$= \sum_{classes} \frac{(\#observed - n \Pr\{data\ in\ class\})^2}{n \Pr\{data\ in\ class\}}$$

In each class, the number (#) of data values "expected" to occur, according to the fitted distribution, is simply the probability of occurrence in that class times the samples size n. The Chi square test is less sensitive to discrepancies in the extreme tails than is the K-S test. The Chi square test operates more naturally for discrete random variables, since to implement it, the range of the data must be divided into discrete classes, or bins.

Among 34 test statistics, 18 cases were best fit using a lognormal distribution, three datasets a Weibull distribution, and 13 datasets a gamma distributions. Thus, the lognormal distribution was selected for the analysis of the $PM_{2.5}$ data in Atlanta. Note that in most cases, test statistics of gamma and Weibull distributions were within the

critical point of the 95 % confidence interval. Thus, gamma and Weibull distributions also fit the $PM_{2.5}$ data well.

Table 2-5. Chi-square test statistics for $PM_{2.5}$ mass

		JST 1999	FTM 1999	TUC 1999	JST 2000	FTM 2000	TUC 2000	JST 2001	FTM 2001	SDK 2001
χ^2 (95% CI)		12.6	12.6	12.6	12.6	12.6	12.6	12.6	12.6	12.6
Lognormal	MM	9.7	12.1	4.2	7.7	3.8	7.8	10.4	1.9	4.9
	MLE	3.9	6.8	1.7	7.7	3.9	7.8	7.9	2.5	4.1
Weibull	MM	9.5	15.8	28.6	36.1	16.7	28.0	18.9	12.7	6.0
	MLE	10.4	14.7	15.0	36.8	14.4	30.6	17.5	13.2	7.0
Gamma	MM	4.9	9.7	4.2	15.8	6.8	16.7	9.9	4.0	2.5
	MLE	4.4	9.0	4.0	16.2	5.9	14.3	10.1	4.1	2.4

		TUC 2001	JST 2002	FTM 2002	SDK 2002	TUC 2002	JST 2003	FTM 2003	SDK 2003
χ^2 (95% CI)		12.6	12.6	12.6	12.6	12.6	12.6	12.6	12.6
Lognormal	MM	9.4	3.7	3.1	3.7	8.4	5.1	2.2	11.3
	MLE	4.3	1.3	3.1	2.5	4.3	7.8	1.4	14.0
Weibull	MM	9.5	14.6	11.6	20.0	5.8	19.5	10.9	30.3
	MLE	10.7	10.0	13.2	31.0	7.8	19.6	11.2	17.8
Gamma	MM	6.2	2.3	3.9	11.6	0.9	8.8	2.5	9.5
	MLE	5.7	1.9	2.8	3.3	1.8	7.5	1.7	9.1

Table 2-6. Kolmogorov-Smirnov (K-S) test statistics for $PM_{2.5}$ mass

		JST 1999	FTM 1999	TUC 1999	JST 2000	FTM 2000	TUC 2000	JST 2001	FTM 2001	SDK 2001
D_n (95% CI)		0.083	0.095	0.084	0.080	0.100	0.092	0.083	0.099	0.087
Lognormal	MM	0.072	0.076	0.041	0.031	0.040	0.035	0.056	0.035	0.029
	MLE	0.056	0.065	0.040	0.031	0.033	0.035	0.058	0.063	0.035
Weibull	MM	0.056	0.100	0.038	0.069	0.064	0.079	0.055	0.054	0.052
	MLE	0.053	0.063	0.040	0.072	0.050	0.075	0.052	0.060	0.053
Gamma	MM	0.059	0.077	0.042	0.044	0.043	0.055	0.035	0.035	0.029
	MLE	0.060	0.069	0.041	0.032	0.034	0.034	0.034	0.042	0.016

		TUC 2001	JST 2002	FTM 2002	SDK 2002	TUC 2002	JST 2003	FTM 2003	SDK 2003
D_n (95% CI)		0.090	0.075	0.092	0.073	0.073	0.074	0.074	0.073
Lognormal	MM	0.055	0.038	0.029	0.029	0.039	0.035	0.017	0.045
	MLE	0.046	0.033	0.032	0.035	0.036	0.046	0.023	0.056
Weibull	MM	0.057	0.036	0.070	0.082	0.063	0.038	0.068	0.035
	MLE	0.043	0.039	0.055	0.088	0.058	0.043	0.062	0.046
Gamma	MM	0.020	0.008	0.031	0.055	0.030	0.032	0.032	0.043
	MLE	0.025	0.022	0.016	0.025	0.022	0.022	0.024	0.026

Three hundred and sixty five random numbers following the lognormal distribution of daily $PM_{2.5}$ mass were generated 1,000 times for each year. The average of the 365 random numbers is considered as the annual mean $PM_{2.5}$ mass concentration. Five thousand annual mean concentrations are calculated for FTM and JST because 1,000 annual mean concentrations for each year from 1999 to 2003 are available. For SDK, 3,000 and for TUC, 4,000 annual mean concentrations are calculated because daily $PM_{2.5}$ data are available only for three years in SDK, and for four years in TUC.

The annual mean levels simulated based on the pdfs of daily $PM_{2.5}$ mass are illustrated in Figure 2-8 (a). The 95th, 50th and 5th percentile of annual mean $PM_{2.5}$ levels are shown. The emission reduction required to decrease the 95th percentile of the yearly $PM_{2.5}$ level below the NAAQS (=15 μg/m^3) is calculated by replacing the yearly mean concentration with the 95th percentile concentration in equation (2-14). The amount of the reduction required is 30.4 % (FTM), 21.8 % (SDK), 42.0 % (TUC), and 32.4 % (JST). As a consequence, additional reductions from 3.4% to 21.3% are necessary when allowance is made for variability. Figure 2-8 (b) represents yearly pollutant levels after the control is applied.

(a) (b)

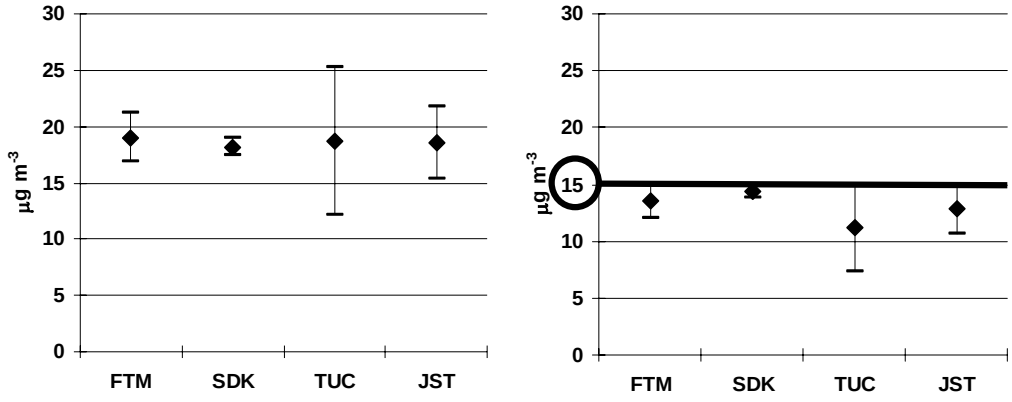


Figure 2-8. Pollutant levels (top: 95th, middle: 50th, and bottom: 5th percentiles). (a) $PM_{2.5}$ concentrations (μg m^{-3}) before control is applied (b) $PM_{2.5}$ concentrations (μg m^{-3}) after control is applied.

Required PM reductions can be estimated by species using these values. Analysis is done for sulfate and OC, the two major species of $PM_{2.5}$ (Figure 2-2). When the reduction in the emissions for $PM_{2.5}$ mass meeting the NAAQS at a 95% CI is given as R (%), the reduction in the emissions for sulfate and OC are calculated as:

$$(R_{oc} * F_{oc} + R_{sulfate} * F_{sulfate}) = R, \quad (2\text{-}15)$$

where, R_{oc}: amount of reduction in OC source and precursor (%),

F_{oc}: fraction of the OC mass to $PM_{2.5}$ mass,

$R_{sulfate}$: amount of reduction in sulfate source (%),

$F_{sulfate}$: fraction of the sulfate mass to $PM_{2.5}$ mass.

Based on the $PM_{2.5}$ species and mass concentrations (Figure 2-2) and the equation (2-14), the amount of reduction in the emission sources and precursors for sulfate and organic are calculated (Figure 2-9).

Reductions (Figure 2-9) are calculated assuming that the change of the amount of emissions for sulfate or OC does not affect other $PM_{2.5}$ species concentrations. However, OC shares emission sources with EC such as wood burning fireplaces and furnaces, and meat cooking combustion process (Hawthorne et al., 1989; Hildemann et al., 1994; Mulhbaier and Williams, 1982). Changes in SO_2 emissions, the major source of sulfate, also affects ammonium and nitrate concentrations because those three species form ammonium nitrate, ammonium sulfate, ammonium bisulfate and so on. Thus, the amount of reduction needs to be modified based on the above information. Here, the information of the correlation between species is used to take into account for the response of other

$PM_{2.5}$ species concentrations to sulfate or OC concentrations. The incorporation of correlation can be achieved by expressing multiple $PM_{2.5}$ species with bivariate joint pdfs.

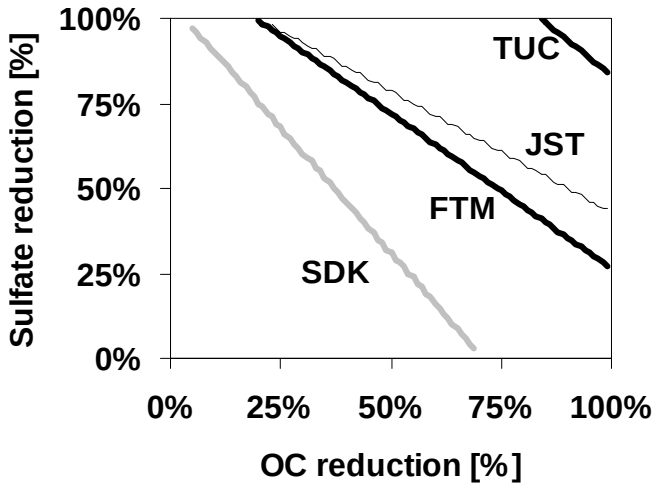


Figure 2-9. Reductions needed for sulfate and OC to meet the NAAQS.

The joint pdf selected is a KS-pdf. Parameters of a KS-pdf are calculated for pairs of OC and EC, and of sulfate and ammonium. SO_2 emission precursors do affect the nitrate concentrations, but the joint pdf is not calculated for nitrate because nitrate concentrations are low in Atlanta (Figure 2-2). Koehler and Symanowski (KS) pdfs conserves marginal distributions. Thus, pairs of sulfate and ammonium, and of organic and EC concentrations from FTM, SDK, TUC, and JST are fitted to univariate distributions, and then the KS-pdf is calculated. For use in this study, KS-pdfs from both

pairs of data are necessary. The KS-pdfs for sulfate and ammonium from FTM and TUC failed the goodness-of-fit test, so only data from SDK and JST are analyzed. The univariate distributions that best fit $PM_{2.5}$ species data are illustrated (Table 2-7). Parameters are estimated using MLE, and the expectations and standard deviation of the distribution and those of the raw data are also shown.

Table 2-7. Univariate distributions that fit $PM_{2.5}$ species data the best

		Correlation	Distribution	Parameters estimated by MLE*		Mean [μg m^{-3}]		Standard deviation [μg m^{-3}]	
						Distribution	Raw data	Distribution	Raw data
SDK	SO_4^{2+}	0.77	lognormal	1.25	0.72	4.5	4.41	3.71	3.31
	NH_4^+		gamma	1.92	0.78	1.5	1.5	1.08	1.13
	OC	0.32	weibull	1.64	6.29	5.63	5.64	3.52	3.52
	EC		weibull	1.29	1.26	1.17	1.17	0.91	0.98
JST	SO_4^{2+}	0.94	lognormal	1.35	0.67	4.84	2.82	3.65	3.38
	NH_4^+		lognormal	0.49	0.59	1.94	1.93	1.26	1.17
	OC	0.82	lognormal	1.34	0.5	4.32	4.32	2.28	2.37
	EC		lognormal	0.24	0.62	1.54	1.53	1.05	1.01

* The two parameters estimated are μ and σ for lognormal distribution, and a and b for weibull and gamma distribution (refer Table 2-2)

Table 2-8. Kolmogorov-Smirnov (K-S) test statistics for $PM_{2.5}$ species

		KS-stat	critical value
SDK	SO_4^{2+}	0.033	0.034
	NH_4^+	0.03	0.034
	OC	0.03	0.035
	EC	0.032	0.035
JST	SO_4^{2+}	0.017	0.031
	NH_4^+	0.014	0.031
	OC	0.016	0.03
	EC	0.013	0.03

Parameters of the KS-pdf estimated based on the univariate distributions are illustrated in Table 2-9, and the goodness-of-fit test results performed by bootstrap re-sampling method are given in Table 2-10. The bootstrap is a resampling method for assessing statistical accuracy (Efron and Tibshirani, 1993). The basic idea is to draw sample with replacement form the original data set where each sample size is the same as the original data set. This is repeated, say K times, rendering K bootstrap samples. Then, the behavior of the K bootstrap samples is examined. In this study, sampling replication is done 30,000 times. The KS-pdf fits well the $PM_{2.5}$ species at 95% CI except for OC—EC pair at JST for xy and xy^2 moments, and for SO_4^{2+}—NH_4^+ pair at SDK for x^2y moment. The deviation of the moments from its sampling is small, and other moments are within the 95% CI for those pairs, so all four pairs illustrated in Table 2-10 are used in this study.

Table 2-9. Parameters of the Koehler-Symanowski (KS) pdf

Species		Parameters		
		α_{11}	α_{12}	α_{22}
JST	SO_4^{2+}, NH_4^+	0.008	0.113	0.02
	OC, EC	0.037	0.31	0.02
SDK	SO_4^{2+}, NH_4^+	0.036	0.078	0.038
	OC, EC	0.029	0.128	0.251

Table 2-10. Goodness of fit tests for cross moments, xy, x^2y, and xy^2 using the boostrap resampling method

Location	SDK		JST	
Species	SO_4^{2+}, NH_4^+	OC, EC	SO_4^{2+}, NH_4^+	OC, EC
M(xy) from the KS-pdf	8.7	7.4	12.3	7.9
2.5th percentile of M(xy)*	8.6	7.2	12.2	8.1
97.5th percentile of M(xy)*	10.4	8.2	13.9	9.2
M(x^2y) from the KS-pdf	23.5	63.8	45.1	49.2
2.5th percentile of M(x^2y)*	25.4	61.6	40.6	55.3
97.5th percentile of M(x^2y)*	39.9	75.9	49.9	72.8
M(xy^2) from the KS-pdf	79.5	14.6	121.7	19.9
2.5th percentile of M(xy^2)*	74.1	14.0	109.1	21.4
97.5th percentile of M(xy^2)*	123.2	19.0	134.9	28.3

* based on the measured data using the boostrap resampling method

The average EC (ammonium) concentration is determined by a conditional KS-pdf. The conditional KS-pdf is the KS-pdf with the condition of the OC(sulfate) of interest. Thus, this method allows the average EC (ammonium) concentrations to be quantified when the amount of the emission sources and precursors of OC (sulfate) changes. The result shows that the amount of reduction needed for sulfate and OC to

meet the standard decreases up to 8% for SDK and 20% for JST (Figure 2-10). The relatively significant reduction in JST compared to that at SDK is due to the higher correlation between species at JST than at SDK (Table 2-7).

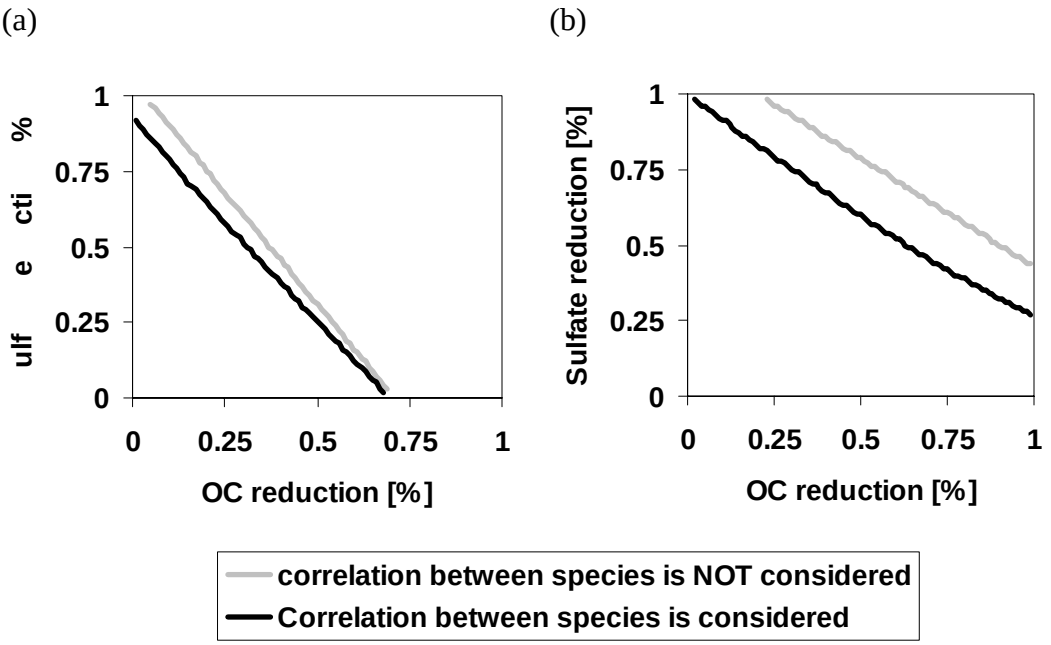


Figure 2-10. Reductions needed for sulfate and OC to meet the NAAQS at 95% CI when the correlation between $PM_{2.5}$ species is taken into account at (a) SDK and (b) JST.

2.4. Conclusions

$PM_{2.5}$ mass measurements in Atlanta, Georgia, show that $PM_{2.5}$ levels exceed the annual NAAQS. The amount of the reduction required to meet the NAAQS in Atlanta is calculated using the rollback method. The linearity assumption, implicit in the rollback equation, is checked via sensitivity analysis using an air quality model. Based on the average of the annual mean $PM_{2.5}$ level, 22% (FTM), 18% (SDK), 21% (TUC), and 20%

(JST) of the emission sources and precursors of $PM_{2.5}$ mass should decrease for annual $PM_{2.5}$ meeting the NAAQS. These results are valid when the annual mean $PM_{2.5}$ does not change temporally. The temporal variation is taken into account using the property that daily $PM_{2.5}$ data follow an underlying distribution. The lognormal distribution was found to be best for the daily $PM_{2.5}$ mass in Atlanta. Three hundred and sixty five random numbers following the lognormal distribution of daily $PM_{2.5}$ mass are generated for 1,000 times for each year. The average of the 365 random numbers is considered as the annual mean $PM_{2.5}$ mass concentration. Based on this analysis, 30% (FTM), 22% (SDK), 42% (TUC), and 32% (JST) of emissions should be decreased for annual $PM_{2.5}$ level to meet the NAAQS.

The analysis was extended for calculating reductions for emissions of each $PM_{2.5}$ species to meet the NAAQS. Because some species have similar emission sources and there are chemical interactions, the reduction of emissions of one species affects the concentrations of other species. In that case, the concentrations of those species are correlated with each other. The correlation between species is taken into account using the Koehler-Symanowski joint pdf, and the amount of the reduction in the emissions can be recalculated. The results show that the reduction of the emission sources and precursors for sulfate and OC meeting the NAAQS decreases up to 8% for SDK and 20% for JST. The relatively significant reduction in JST compared with that in SDK is due to the higher correlation between species in JST than that in SDK. The results in these analyses can be a useful reference in developing a $PM_{2.5}$ control strategy in the Atlanta area.

Acknowledgement

This research was supported by the U.S. Environmental Protection Agency under agreements RD82897602, RD83107601, and RD83096001.

References

Baumann, K., Chang, M. E., Russell, A. G., and Edgerton, E. S. (2005). Seasonal Variability of Organic Mass Contribution to PM2.5 within Metro Atlanta are Further Downwind. *AAAR international specialty conference -- Particulate Matter Supersites Program and Related Studies, Atlanta, Georgia*.

Butler, A. J. (2000). Temporal and spatial analysis of PM2.5 mass and composition in Atlanta. *Ph. D. Thesis, School of Civil and Environmental Engineering, Georgia Institute of Technology*.

Butler, A. J., Andrew, M. S., and Russell, A. G. (2003). Daily sampling of PM2.5 in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. *Journal Of Geophysical Research-Atmospheres* **108**.

Chock, D. P., and Sluchak, P. S. (1986). Estimating Extreme Values Of Air-Quality Data Using Different Fitted Distributions. *Atmospheric Environment* **20**, 989-993.

deNevers, N., Neligan, R. E., and Slater, H. H. (1977). Air Pollution. *Academic Press, New York* **vol. 5**.

Efron, B. and Tibshirani, R.J. (1993). An Introduction to Bootstrap, *Chapman and Hall*.

Fletcher, R. (2003). Practical methods of optimization. *Chichester, Wiley*.

Genikhovich, E. L., Ziv, A. D., Iakovleva, E. A., Palmgren, F., and Berkowicz, R. (2005). Joint analysis of air pollution in street canyons in St. Petersburg and Copenhagen. *Atmospheric Environment* **39**, 2747-2757.

Georgopoulos, P. G., and Seinfeld, J. H. (1982). Statistical Distributions Of Air Pollutant Concentrations. *Environmental Science & Technology* **16**, A401-&.

Hadley, A., and Toumi, R. (2003). Assessing changes to the probability distribution of sulphur dioxide in the UK using a lognormal model. *Atmospheric Environment* **37**, 1461-1474.

Hansen, D. A., Edgerton, E. S., Hartsell, B. E., Jansen, J. J., Kandasamy, N., Hidy, G. M., and Blanchard, C. L. (2003). The southeastern aerosol research and characterization study: Part 1-overview. *Journal Of The Air & Waste Management Association* **53**, 1460-1471.

Hawthorne, S. B., Krieger, M. S., Miller, D. J., and Mathiason, M. B. (1989). Collection And Quantitation Of Methoxylated Phenol Tracers For Atmospheric-Pollution From Residential Wood Stoves. *Environmental Science & Technology* **23**, 470-475.

Hildemann, L. M., Klinedinst, D. B., Klouda, G. A., Currie, L. A., and Cass, G. R. (1994). Sources Of Urban Contemporary Carbon Aerosol. *Environmental Science & Technology* **28**, 1565-1576.

Holland, D. M., and Fitzsimons, T. (1982). Fitting Statistical Distributions To Air-Quality Data By The Maximum-Likelihood Method. *Atmospheric Environment* **16**, 1071-1076.

Kao, A. S., and Friedlander, S. K. (1995). Frequency distribution of PM10 chemical components and their sources. *Environmental Science & Technology* **29**, 19-28.

Koehler, K. J., and Symanowski, J. T. (1995). Constructing multivariate distributions with specific marginal distributions. *Journal of multivariate analysis* **55**, 261-282.

Kruize, H., Hanninen, O., Breugelmans, O., Lebret, E., and Jantunen, M. (2003). Description and demonstration of the EXPOLIS simulation model: Two examples of modeling population exposure to particulate matter. *Journal Of Exposure Analysis And Environmental Epidemiology* **13**, 87-99.

Lu, H. C. (2002). The statistical characters of PM10 concentration in Taiwan area. *Atmospheric Environment* **36**, 491-502.

Lu, H. C., and Fang, G. C. (2002). Estimating the frequency distributions of PM10 and PM2.5 by the statistics of wind speed at Sha-Lu, Taiwan. *Science Of The Total Environment* **298**, 119-130.

Manomaiphiboon, K., and Russell, A. G. (2003). Formulation of joint probability density function of velocity for turbulent flows: an alternative approach. *Atmospheric Environment* **37**, 4917-4925.

Mulhbaier, J. L., and Williams, R. L. (1982). Fireplaces, furnaces, and vehicles as emission sources of particulate carbons. *Particulate Carbon-Atmospheric Life Cycle, edited by G. T. Wolff*, 185-205.

Ott, W., Wallace, L., and Mage, D. (2000). Predicting particulate (PM10) personal exposure distributions using a random component superposition statistical model. *Journal Of The Air & Waste Management Association* **50**, 1390-1406.

Park, S.-K., Cobb, C. E., Wade, K., Mulholland, J., Hu, Y., and Russell, A. G. (Submitted). Uncertainty in Air Quality Model Evaluation from Spatial Variation. *Atmospheric Environment*.

Park, S.-K., and Russell, A. G. (2003). Sensitivity of PM 2.5 to emissions in the Southeast. *Proceedings of the Models-3 User's workshop, Research Triangle Park, North Carolina.*

Rumburg, B., Alldredge, R., and Claiborn, C. (2001). Statistical distributions of particulate matter and the error associated with sampling frequency. *Atmospheric Environment* **35**, 2907-2920.

Singh, K. P., Warsono, Bartolucci, A. A., and Bae, B. (2001). Mathematical modeling of environmental data. *Mathematical And Computer Modelling* **33**, 793-800.

US-EPA (2004a). Air Quality Criteria for Particulate Matter (October 2004). *EPA 600/P-99/002aF-bF*.

US-EPA (2004b). SMOKE User's manual. *http://cf.unc.edu/cep/empd/products/smoke*.

Wark, K., Warner, C. F., and Davis, W. T. (1997). Air pollution. Its origin and control. *Prentice Hall*, 560p.

CHAPTER 3

UNCERTAINTY IN AIR QUALITY MODEL EVALUATION FOR PARTICULATE MATTER DUE TO SPATIAL VARIATIONS IN POLLUTANT CONCENTRATIONS

(S.-K. Park, C.E. Cobb, K. Wade, J. Mulholland, Y. Hu, A.G. Russell.

Atmospheric Environment, in press)

Abstract

Air quality model performance is usually evaluated by examining the relative agreement between volume-averaged simulations and point measurements as volume-averaged measurements are seldom available. Because the two values have a different spatial scale, accurate model evaluation is complicated by this mismatch in areas when the pollutant gradient is large. Uncertainty in the air quality model evaluation from the spatial variability of $PM_{2.5}$ is quantitatively examined, and how much of model error might be explained by such variability is calculated. Added uncertainty of model performance is analyzed by comparing performance metrics between simulated concentrations and observations at one station between simulated levels and interpolated fields from observations. Normalized differences of the performance metrics (e.g., mean fractional error; MFE) calculated in these two ways indicate the uncertainty of the model performance due to spatial variation. Normalized difference of MFE for $PM_{2.5}$ mass is around 17% in July 2001 and 15% in January 2002. To decrease the uncertainty, it has been suggested that observations be used only from spatially representative stations.

When model performance is calculated with data from spatially representative stations, uncertainty decreased, and overall model performance improves. For example, MFE is seen to decrease up to 14% for $PM_{2.5}$ mass and species concentrations, suggesting that up to 14% of MFE can be explained by the spatial variability of $PM_{2.5}$. These results indicate that comparison between observed and simulated concentrations should not be used alone to assess performance of air quality models. Also, spatial variability should be considered in setting model performance goals.

Key words: Model performance; Spatial variability; Representative stations

3.1. Introduction

Grid-based photochemical air quality models are essential tools for use in air quality management and scientific investigation. Using these models for such applications with confidence necessitates thorough model evaluation, which traditionally is achieved by comparing simulated concentrations with those observed. However, such a performance assessment has limitations that arise from the modeling method and observations. Ideally, predicted concentrations should be compared with volume-averaged measurements, but volume-averaged measured concentrations are seldom available. Monitoring stations cannot be mounted densely in wide regions due to cost, and there are few techniques that directly provide volume-averaged concentrations. Thus, simulated concentrations are usually compared with concentrations measured at specific monitoring locations, with each monitor representing a point (in space) measurement. Concentrations measured at monitoring sites can differ substantially from average concentrations in the area if pollutant concentration gradients are high. Therefore, model performance evaluated by comparing between point observations and volume-averaged

simulations may not represent how well the model actually simulates air pollution dynamics.

The implication of the spatial inhomogeneity on air quality model performance has been investigated by quantifying the spatial variability for ozone (O_3), carbon monoxide (CO), and nitrogen oxide (NO) from the difference between the measurement at the monitor and the interpolated concentrations from other monitors (McNair et al., 1996). Because the amount of spatial variability was similar to that of the air quality model error, the authors concluded that spatial variability in observed pollutant concentrations should be taken into account in developing model performance guidelines. In a separate study, a quantitative measure for the spatial representativeness of ground level ozone concentrations is analyzed using the hourly ozone concentrations at 300 monitors in Germany as a means to compare modeled and measured data (Tilmes and Zimmermann, 1998). This analysis suggested that a radius of representativness is about 4 km for ozone. The evaluation of models with grid sizes larger than 4 km may have substantial error due to spatial variation. While there is an awareness that model performance depends on spatial variability, there has been no effort to analyze the dependency quantitatively.

Here, observations and air quality model results are used to quantify spatial variability in particulate matter and how that adds uncertainty in model evaluation for both individual $PM_{2.5}$ species and total mass. In this paper, the amount of calculated model error "error" that might be explained by such variability at the regional scale is investigated, and ways to reduce added uncertainty in model performance evaluation are proposed.

3.2. Methods

Pollutant concentrations studied for spatial analysis and model evaluation include daily $PM_{2.5}$ mass and various species (sulfate, nitrate, ammonium, elemental carbon, and organic carbon) with particular focus on the Atlanta area for two years (2002-2003) and over the continental United States area for July 2001 and January 2002. The latter two one-month periods have additional observations as part of the Eastern Supersite Program (ESP 01/02). Monitoring data were obtained from the Assessment of Spatial Aerosol Composition in Atlanta (ASACA) project (Butler et al., 2003), the Southeastern Aerosol Research and Characterization (SEARCH) study (Hansen et al., 2003), the Environmental Protection Agency's Speciated Trends Network (EPA-STN) databases (Jang et al., 2004), and the Interagency Monitoring of Protected Visual Environments (IMPROVE) network (Ames and Malm, 2001) (Figure 3-1).

EPA's Models-3 was applied over a domain covering the United States using the unified Regional Planning Organization (RPO) national grid with a 36km resolution (Figure 3-2). Models-3 used includes the Community Multi-scale Air Quality (CMAQ v4.3) model (Byun and Ching, 1999), the Sparse Matrix Operator Kernel Emissions (SMOKE v1.5) (US-EPA, 2004d), and NCAR's 5th generation Mesoscale Model (MM5 v3.5.3) for meteorological modeling (NCAR, 2003) (see Table 3-1 for details). Meteorological fields were evaluated with the Barnes objective analysis scheme (Koch et al., 1983) using the TDL surface hourly data (UCAR, 2003e), which were not used in the four-dimensional data assimilation. Mean errors (MEs) in temperature, specific humidity, and wind speed were 1.7 ^{o}C / 2.1 ^{o}C (July, 2001 / January, 2002), 1.8g kg^{-1} / 0.5g kg^{-1},

and 1.3m sec^{-1} / 1.4m sec^{-1}, respectively (Park and Russell, 2003a). These values are within the benchmarks for the metrological model evaluation (Emery et al., 2001).

Table 3-1. Detailed Information of the air quality modeling system

MODEL	INDEX		COMMENTS
CMAQ	Chemistry		SAPRC99 (Carter, 2000)
	Chemistry solver		Modified Euler Backward Iterative (MEBI)
	Aerosol equilibrium		ISORROPIA (Nenes, 1996; Nenes et al., 1998)
	Aerosol dynamics		AERO3
	Deposition velocity		AERO_DEPV2
	Advection (horizontal and vertical)		Piecewise Parabolic Method (PPM)
	Cloud processing		Regional Acid Deposition Model (RADM)
SMOKE	Anthropogenic emissions	Inventory	Survey from Fall Line Air Quality (FAQS) project for GA (Hu et al., 2003; Unal et al., 2003)
			1999 NEI final v2.0 for other states (US-EPA, 2003a)
			(0.25 was multiplied to fugitive dust emissions) (Placet et al., 2000a)
		Growth factor	EGAS 4.0 (US-EPA, 2004c)
		Hourly emissions	Continuous Emission Monitoring (CEM) for major point sources (NO_x, SO_2,) (US-EPA, 2003b)
		Spatial surrogate	Based on the 2000 census data (US-EPA, 2003c)
	Biogenic emissions		Biogenic Emission Land cover Database v3 (BELD3) (US-EPA, 2004b)
MM5	Microphysics		Simple ice microphysics
	Cumulus scheme		Kain-Fritsch
	Boundary layer		Pleim-Chang
	Radiation scheme		Rapid Radiative Transfer Model (RRTM)
	Land surface model		Pleim-Xiu
	Data for four dimensional data assimilation		NCEP ETA model outputs for the GCIP project (UCAR, 2003a)
			NCEP ADP observation data (UCAR, 2003b, 2003c, 2003d)

(a)

(b)

(c)

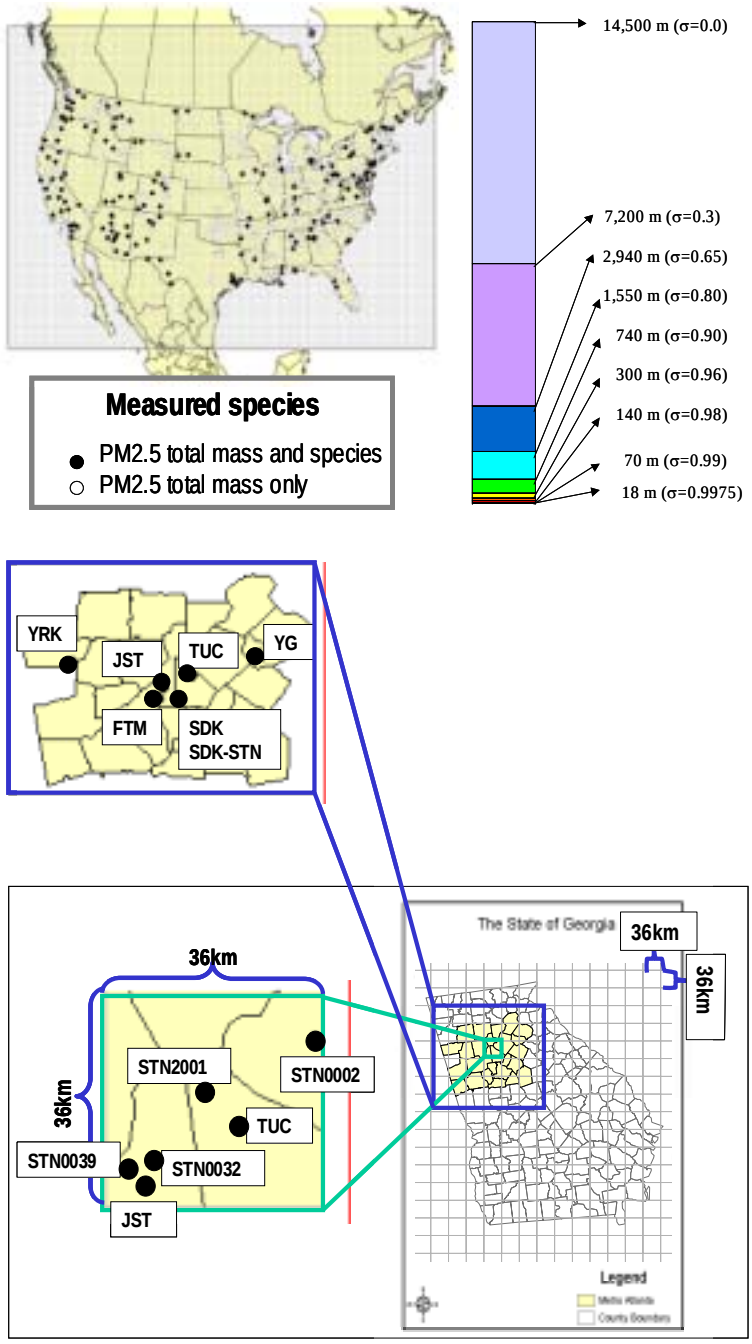


Figure 3-1. (a) Horizontal and vertical structures of the air quality model domain, and $PM_{2.5}$ species and mass monitors in the United States. (b) $PM_{2.5}$ species and mass monitors in Atlanta. (c) $PM_{2.5}$ mass monitors in one grid of the air quality model.

Concentrations from one monitor are compared with the concentrations in the area surrounding the monitor to calculate the spatial inhomogeneity. In the Atlanta area, $PM_{2.5}$ concentrations in the surrounding area were calculated using the average of $PM_{2.5}$ concentrations from other monitors located within 120 km. In the United States, some locations have less than four monitors within 120 km, so monitors within 180 km were used to calculate concentrations in the surrounding area. Average concentrations were calculated by an inverse-distance-squared interpolation (McNair et al., 1996):

$$I(m) = \frac{\sum_{i=1,\ i\neq m}^{N} O_i * W_i}{\sum_{i=1,\ i\neq m}^{N} W_i} \quad , \tag{1}$$

where *I(m)* is the interpolated pollutant concentration for station *m*, *N* is the number of monitoring stations, O_i is the observed pollutant concentration at station *i*, and W_i is the weight, calculated as;

$$W_i = \begin{cases} \frac{1}{r_i^2} & if \ \ r_i \leq 180km \\ 0 & if \ \ r_i > 180km \end{cases}, \tag{2}$$

and r_i is the distance from station *m* to station *i*. Stations, less than 20% of the total, which do not have other stations within 180 km of each other were not included in this analysis.

Spatial inhomogeneity was quantified by applying traditional model performance metrics, i.e., the concentrations from one monitor were compared with the observed concentrations in the area surrounding the monitor. Calculated performance metrics include the Mean Bias (MB), Mean Error (ME), Normalized Mean Bias (NMB),

Normalized Mean Error (NME), Mean Fractional Bias (MFB), and Mean Fractional Error (MFE) (Boylan et al., 2005). Carbon concentrations were measured by two different methods, thermal optical transmittance (TOT) and thermal optical reflectance (TOR). Sites from ASACA and STN used TOT, and those from IMPROVE and SEARCH used TOR. Thus, the surrounding concentrations for organic and elemental carbon concentrations were calculated separately.

Uncertainty in the model performance introduced by spatial variation was quantified over the United States by comparing performance metrics between simulated concentrations and observations at one station with that between simulated levels and interpolated fields from observations at surrounding stations. The difference of the performance metrics calculated in the two different ways at each station corresponds to the uncertainty of the model performance due to spatial variation for that station.

3.3. Results and Discussion

3.3.1. Spatial variability of particulate matter

Spatial variability was calculated using performance metrics that compare observed concentrations at each monitor with observed concentrations in the surrounding area of the monitor. The spatial variability was calculated separately for the Atlanta area using data from 2002 to 2003 (Table 3-2), and for the United States using data for July 2001 and January 2002 (Table 3-3 and Figure 3-2a). Because the average observation (for Atlanta) and the interpolated concentration (for the United States) were derived from observations in the surrounding monitors, the bias measures (MB, NMB, MFB) are very small, and do not have important implications. Spatial variability and the possible impacts on calculated model performance can be judged based on the error metrics (ME,

NME, MFE). Spatial variability of $PM_{2.5}$ species leads to MFEs of 30% to 59% in Atlanta (Table 3-2), and MFEs of 28% to 84% in the United States for different species (Table 3-3 and Figure 3-2a).

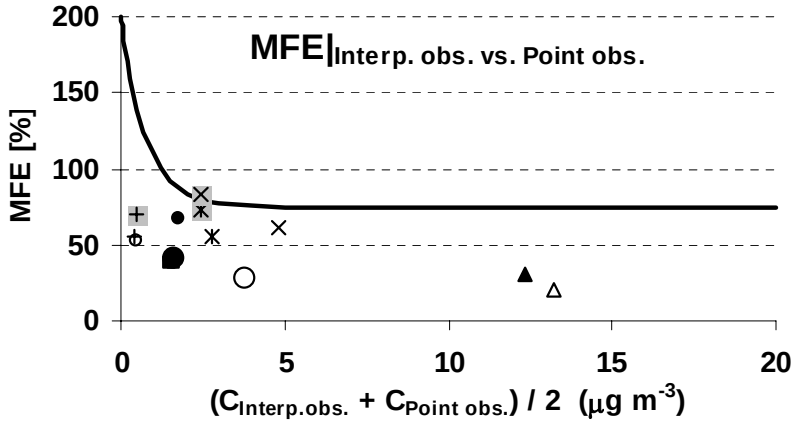
Table 3-2. Spatial variability of $PM_{2.5}$ mass and species in the Atlanta area from January 2002 to December 2003.

Species	Number of obs.	Mean conc. ($\mu g/m^3$)	MB ($\mu g/m^3$)	ME ($\mu g/m^3$)	NMB (%)	NME (%)	MFB (%)	MFE (%)
PM2.5 total mass	2903	17.50	0.09	2.20	0.5	12.6	2.0	13.2
Sulfate	2781	4.10	-0.02	1.23	-0.6	30.0	6.2	31.4
Nitrate	2732	0.91	-0.01	0.34	-0.9	37.5	8.7	38.7
Ammonium	2611	1.52	-0.02	0.50	-1.4	33.2	6.5	37.9
Organic Carbon	2045	4.70	0.00	1.83	0.0	38.9	2.9	41.9
Elemental Carbon	1890	0.81	0.00	0.47	0.0	58.0	4.1	59.4
Total Carbon	2349	5.65	-0.06	1.54	-1.0	27.3	4.3	29.5
Soil Dust	1758	5.57	0.26	2.47	4.7	44.3	13.2	46.3

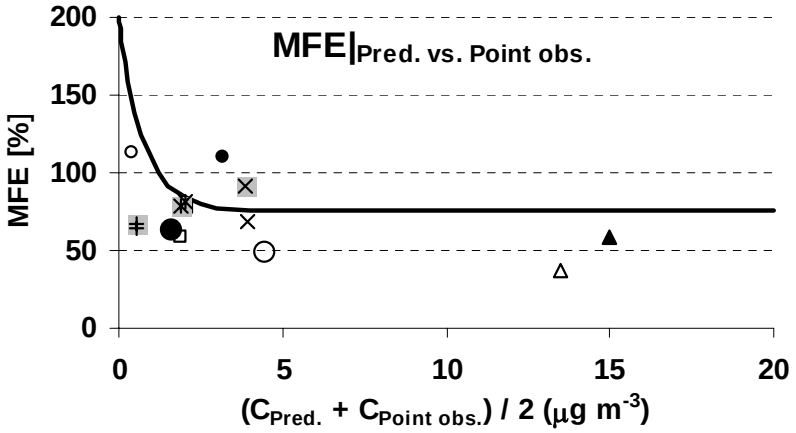
Table 3-3. Spatial variability of $PM_{2.5}$ mass and species in the United States calculated using data in all stations.

Species	Period	Number of sites	Number of obs.	Mean obs. conc. ($\mu g/m^3$)	MB ($\mu g/m^3$)	ME ($\mu g/m^3$)	NMB (%)	NME (%)	MFB (%)	MFE (%)
PM2.5 total mass	Jul. 2001	1206	14798	13.22	0.28	2.41	2.1	18.2	5.3	20.4
	Jan. 2002	1153	14160	12.35	0.48	3.26	3.9	26.4	9.6	30.0
Sulfate	Jul. 2001	189	1855	3.81	0.05	0.88	1.4	23.0	4.7	28.1
	Jan. 2002	189	1695	1.60	0.06	0.43	3.4	27.1	11.0	40.8
Nitrate	Jul. 2001	176	1630	0.48	0.03	0.32	5.5	67.2	12.7	52.1
	Jan. 2002	171	1501	1.78	0.12	1.09	6.8	61.2	17.5	66.8
Ammonium	Jul. 2001	69	863	1.64	0.00	0.58	0.3	35.5	5.4	37.3
	Jan. 2002	64	654	1.49	0.03	0.47	2.0	31.5	5.1	38.3
Organic carbon	Jul. 2001	189	1910	2.79	0.17	1.56	6.0	55.8	12.3	55.0
	Jan. 2002	183	1675	2.44	0.27	1.76	11.2	72.2	22.3	72.5
Elemental carbon	Jul. 2001	189	1907	0.43	0.04	0.25	10.2	58.0	14.6	55.6
	Jan. 2002	183	1660	0.50	0.05	0.38	10.5	76.6	19.1	69.8
Soil dust	Jul. 2001	59	497	4.81	0.09	2.92	1.9	60.6	6.1	61.4
	Jan. 2002	55	386	2.44	0.03	1.97	1.2	80.8	5.2	83.6

(a)



(b)



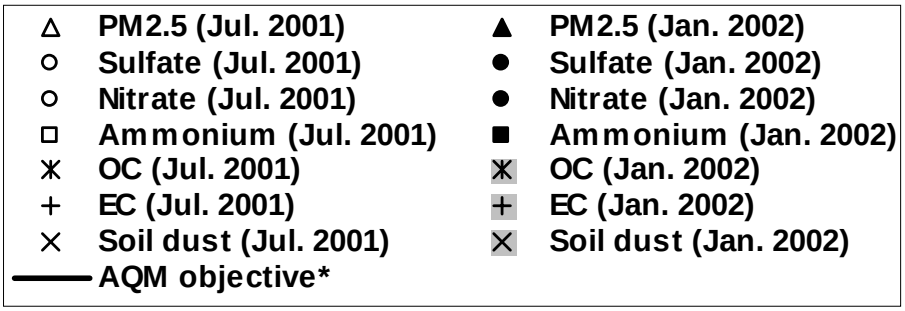


Figure 3-2. Mean fractional error (MFE) of simulated and interpolated concentrations against observed concentrations in the United States, and suggested objectives (Boylan, 2004) based on analysis of modeling studies.
*AQM objective (Boylan, 2004):

$$MFE\ [\%] \le 125e^{\frac{-2(C^o+C^e)}{3[\mu g/m^3]}} + 75$$

where C^e is average model-estimated/interpolated concentration, C^o is average observed concentration.

Spatial variability was higher for primary pollutants (e.g., elemental carbon and soil dust) and lower for secondary pollutants (e.g., sulfate, nitrate, and ammonium), in general. Spatial variability in different species can be explained by particular emissions and formation characteristics. Sulfate had the lowest spatial variability among the $PM_{2.5}$ species. Sulfate is primarily formed either from the gas-to-particle conversion of SO_2 in the atmosphere or from reactions in the aqueous phase. Because of the slow rates of formation and removal, sulfate becomes relatively well mixed (Roberts and Friedlander, 1980). Nitrate is also a secondary pollutant produced by the oxidation of NO_x ($NO+NO_2$). The oxidation rate of NO_x ranges from 5 to 50% per hour (Spicer et al., 1981), faster than that of SO_2. Also, nitric acid deposits rapidly. Thus, nitrate is spatially less homogeneously distributed than sulfate. Ammonium is also a secondary pollutant formed as ammonia (NH_3) neutralizes H_2SO_4 and HNO_3. The amount of ammonium is highly dependent on the relative amounts of H_2SO_4 and NH_3. Spatial variability of ammonium is also relatively small, usually being dominated by its association with sulfate. Soil dust is a direct emission and settles relatively rapidly, so spatial variability of soil dust and its associated elements is very high. Because elemental carbon is a primary pollutant, the spatial variability is relatively high. Major sources of elemental carbon include diesel engines, particularly heavy-duty trucks, wood burning fireplaces and furnaces, and meat cooking combustion process (Gray and Cass, 1998). Organic carbon is emitted directly or formed from the condensation of low-volatility hydrocarbons. Thus, the spatial variability of organic carbon is between that of purely primary and secondary pollutants. Major primary sources of organic carbon include diesel and gasoline-burning engines, wood

burning, meat cooking operations, some industrial processes, and biogenic sources (Zheng et al., 2002).

3.3.2. Uncertainty of Air Quality Model Performance due to Spatial Variation

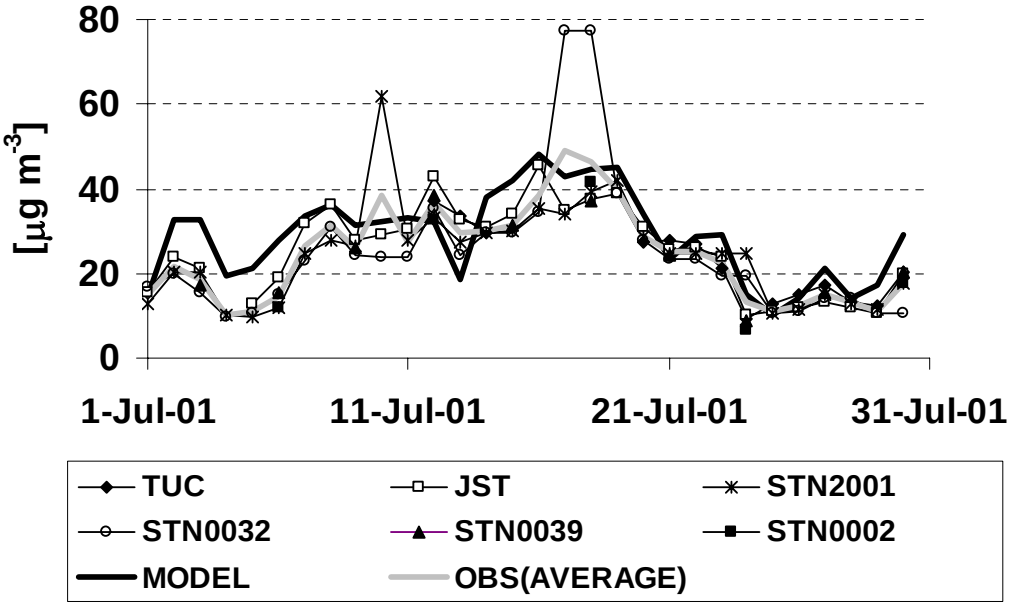
The air quality model was evaluated by comparing the volume-averaged simulated concentrations with the point measurements in the United States (Table 3-4) using six performance metrics (Boylan et al., 2005). In addition to performance metrics used to assess spatial variability, MB, NMB, and MFB are included and, have important implications. In particular, the mean fractional error (MFE) was compared with proposed air quality model objectives as MFE is considered a key indicator (Boylan et al., 2005). All species except nitrate and those associated with soil dust in January 2002 met the objectives (Figure 3-2b). Nitrate is difficult to model due to its volatility and its sensitivity to temperature, relative humidity, and ammonia availability (Russell et al., 1983). Soil dust emissions are uncertain (Pace, 2003). Model performance was better for secondary species, such as sulfate and ammonium, than it was for primary species, such as elemental carbon and soil dust. Similar to the previous study (McNair et al., 1996), model performance is compared to spatial variability (Figures 3-3a and 3-3b). Not surprisingly, MFE between simulated and observed concentrations was larger than MFE of interpolated and observed concentrations, though for some species, the difference is remarkably small. This result also suggests that the added uncertainty in the model error from the spatial inhomogeneity should not be ignored. This issue is further examined below.

Table 3-4. Overall air quality model performance for $PM_{2.5}$ mass and species in the United States calculated using data in all stations.

Species	Period	MB ($\mu g/m^3$)	ME ($\mu g/m^3$)	NMB (%)	NME (%)	MFB (%)	MFE (%)
PM2.5 total mass	Jul. 2001	0.50	4.70	3.8	35.5	2.7	37.6
	Jan. 2002	5.24	8.80	42.4	71.2	26.2	58.1
Sulfate	Jul. 2001	1.34	1.95	35.2	51.4	23.2	49.0
	Jan. 2002	0.12	0.86	7.2	53.8	15.8	62.6
Nitrate	Jul. 2001	-0.11	0.39	-22.0	81.9	-80.3	112.1
	Jan. 2002	2.87	3.48	160.8	195.1	74.6	109.8
Ammonium	Jul. 2001	0.53	0.93	32.4	56.5	31.4	58.0
	Jan. 2002	1.27	1.69	85.2	113.0	51.1	77.5
Organic carbon	Jul. 2001	-1.50	1.81	-53.6	64.7	-58.0	81.4
	Jan. 2002	-1.12	1.54	-45.8	63.1	-30.6	78.7
Elemental carbon	Jul. 2001	0.19	0.38	43.0	87.2	9.1	63.6
	Jan. 2002	0.14	0.40	26.9	80.6	-1.9	67.8
Soil dust	Jul. 2001	-1.81	2.91	-37.5	60.4	-26.2	68.5
	Jan. 2002	2.90	3.67	118.8	150.6	68.4	91.8

One of the model grids located over the Atlanta area contains six $PM_{2.5}$ mass monitors: Fire Station 8 (STN0039), Jefferson Street (JST), E. Rivers School (STN0032), Doraville health center (STN2001), Tucker (TUC), Gwinnett Tech (STN0002) (Figure 3-1c). Although these six monitors are located in the same grid, they have different daily $PM_{2.5}$ total masses showing that significant spatial variability of pollutant concentrations exists within 36km (Figure 3-3a). Thus, air quality model performance differed when observations from different monitors are used. Also, the model performance calculated using the observation at the monitor is markedly different from that using the average observation. These results suggest that the air quality model performance evaluated using point measurement includes uncertainty due to pollutant gradients (Figure 3-3b).

(a)



(b)

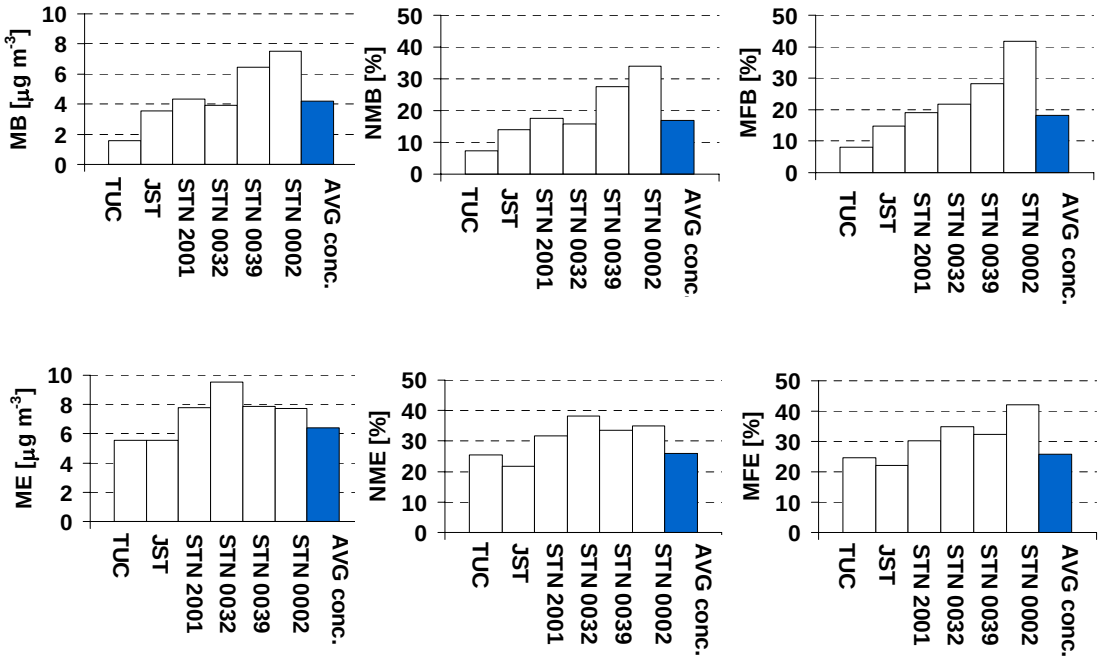


Figure 3-3. (a) Simulated and observed $PM_{2.5}$ mass. (b) Air quality model performance of the same grid based on multiple observations.

Uncertainty of the model performance from spatial variation was investigated over the United States by comparing performance metrics between simulated concentrations and observations at one station with those between simulated levels and interpolated fields from observations. The normalized difference of the performance metrics calculated in two different ways at each station corresponds to the added uncertainty of the model performance due to spatial variation for that station. Median, 25^{th} and 75^{th} percentiles of the normalized difference from the two performance metrics are plotted in Figure 3-4. By representing 25^{th} and 75^{th} percentiles, effects of the varying types of stations are removed to a large degree (Tilmes and Zimmermann, 1998). For example, the median of the normalized difference of MFE for $PM_{2.5}$ mass concentrations is 17% in July 2001 and 15% in January 2002 (Figure 3-4). Interpolation, however, can introduce and increase discrepancies. Simulated values can account for local emissions and $PM_{2.5}$ dynamics than are found in interpolated values. In addition, the interpolation technique used in this study may not always be the optimal approach for accurately deriving the volume-averaged of the observed values. Studies suggest that the interpolated value may depend on the interpolation technique, although the difference typically is not significant (Jarvis and Stuart, 2001).

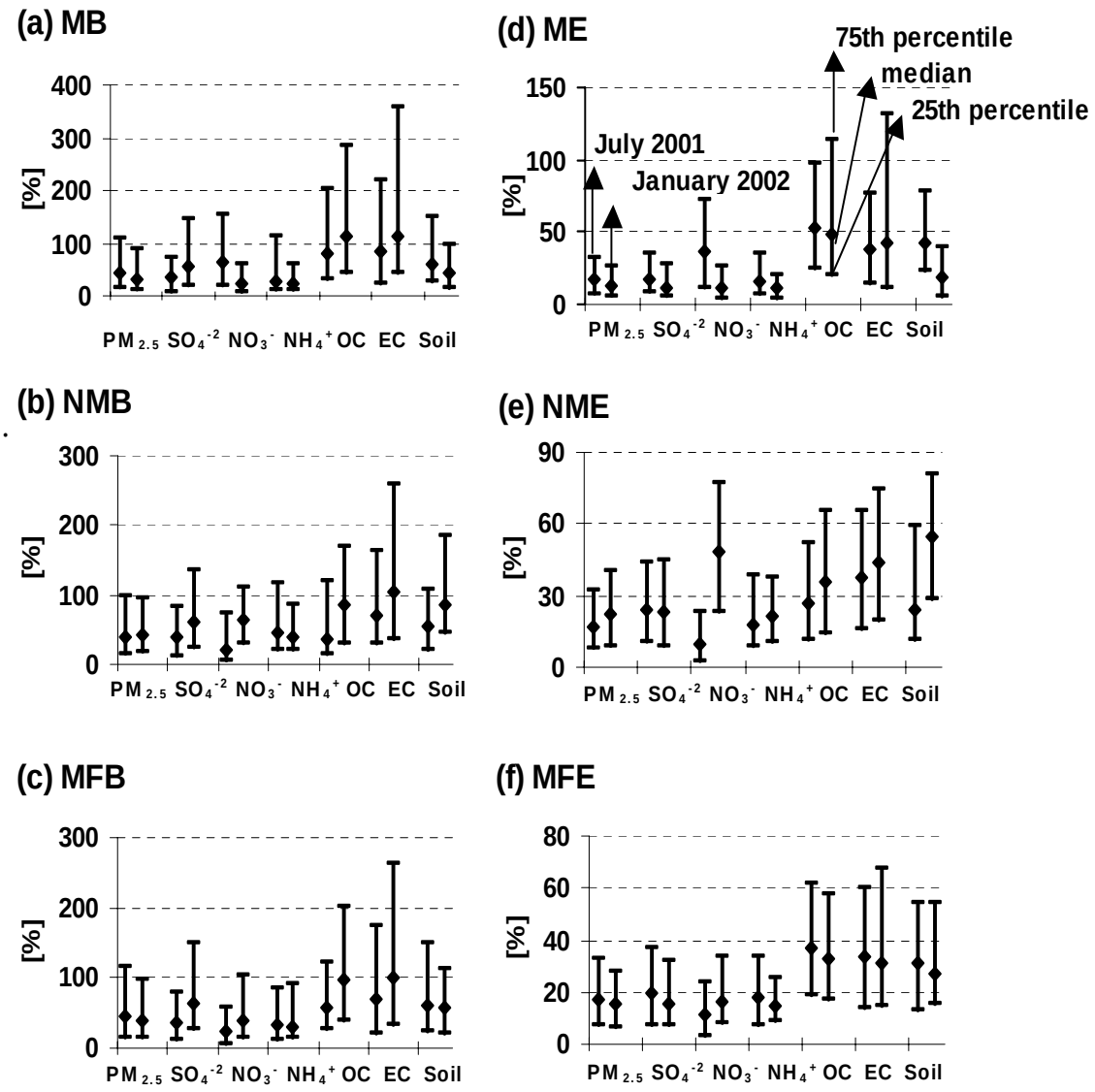


Figure 3-4. Normalized differences* in performance metrics between simulations and observations from that between simulated levels and interpolated fields for each station in the United States.

*Normalized difference of MB:

$$\left| \frac{\textit{prediction vs. interpolation} - \mathrm{MB}|_{\textit{prediction vs. observation}}}{MB|_{\textit{prediction vs. observation}}} \right| \times 100\ [\%]$$

One possible approach to get a better assessment of model performance is to use measurements from spatially representative stations. Representative sites for $PM_{2.5}$ were defined to be those with less than a 20% variation as compared to the annual mean $PM_{2.5}$ concentration of the area (Watson et al., 1997). In this study, sites were judged to be representative only if the monthly mean $PM_{2.5}$ concentration of the site deviates less than 20% from that calculated by interpolating pollutant concentrations from other stations located within 180 km (equations (1) and (2)). From 20% to 30% of the total stations were found to be representative for primary species, and from 50% to 70% for secondary species. As expected, the number of representative stations for primary species was much smaller than that for secondary species.

Uncertainty in model performance from spatial variation decreased when only spatially representative stations are used to calculate model performance. For example, the median of normalized differences of MFE for $PM_{2.5}$ mass concentration calculated with two different methods decreased from 17% to 14% in July 2001 and from 15% to 10% in January 2002. In addition, overall model performance of $PM_{2.5}$ from spatially representative stations improved when compared with the performance of $PM_{2.5}$ from all stations. Among performance metrics, MFE decreased up to 14%; however, it actually increased in two cases: organic carbon and soil dust in January 2002 (Figure 3-5). Soil dust was poorly simulated in this period, further suggesting that more basic issues are involved (e.g., an inaccurate inventory). In a separate investigation (Park et al., submitted for publication), it was found that soil dust levels increase with wind, which is not surprising. On the other hand, the inventory remains constant and higher winds tend to dilute the emissions more leading to lower simulated concentrations.

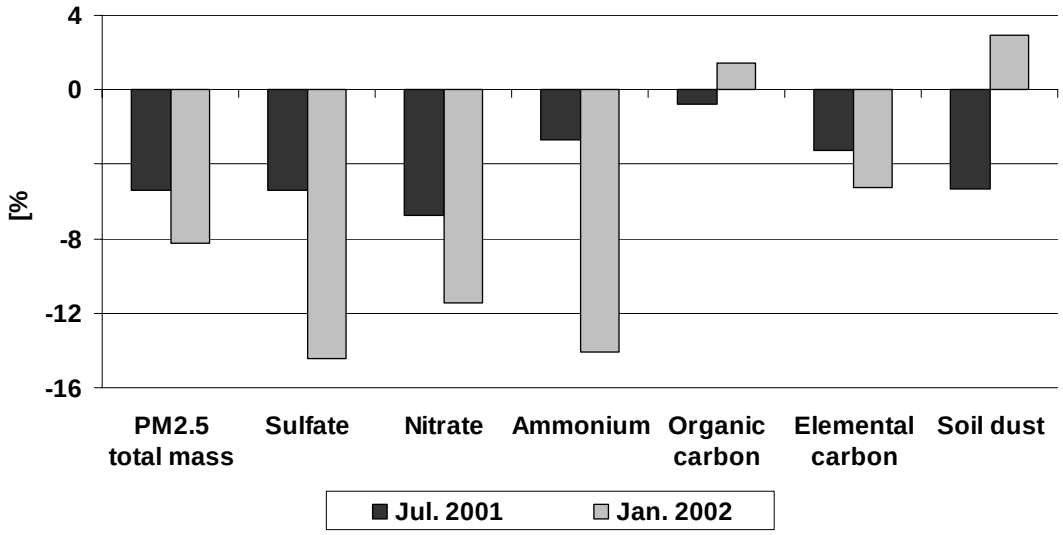


Figure 3-5. Normalized change* of mean fractional error between simulations and observations for all stations from that between simulations and observations for representative stations.
*Normalized change of MFE:

$$\frac{MFE\,|_{representative\ \ stations} - MFE\,|_{all\ \ stations}}{MFE\,|_{all\ \ stations}} \times 100\ [\%]$$

The analyses performed in this paper suggest that spatial variability contributes to uncertainty in model performance. However, spatial variability affected the overall model performance only a moderate amount, implying that there are other sources of the model error, such as emission inventory, meteorology, chemical mechanism parameters, and numerical routine that contribute the model error significantly. Further investigation of these sources of uncertainty should be conducted in the future.

3.4. Conclusions

Air quality model performance is determined by the relative agreement between observed and simulated concentrations. Models predict volume-averaged concentrations,

whereas monitors measure concentrations at a single point in space. This introduces uncertainty in model performance evaluation if pollutant concentrations are spatially inhomogeneous. Spatial variability of PM2.5 mass and species concentrations assessed by comparing interpolated observations to point observations led to calculated performance metrics comparable to model error in magnitude, suggesting that spatial variability impacts model performance. Model performance degradation due to spatial variation in PM2.5 is quantified by comparing model performance using interpolated observations with model performance using point observations. For example, the median of normalized differences of MFE for $PM_{2.5}$ mass concentrations is 17% in July 2001 and 15% in January 2002. When spatially representative stations are used, the median of normalized differences of MFE for $PM_{2.5}$ mass concentrations is only 15% in July 2001 and 10% in January 2002. Overall MFE from representative stations generally improves (up to 14%). Therefore, this study suggests that up to 14% of MFE for $PM_{2.5}$ species and mass concentrations are due to spatial variability in $PM_{2.5}$. The analysis performed in this paper suggests that spatial variability degrades model performance moderately. However, spatial inhomogeneity does not appear to be the major contributor to model error. Additional study for other sources of error should be performed in the future.

Acknowledgement

This research was supported by the U.S. Environmental Protection Agency under Agreements RD82897602, RD83107601, and RD83096001. We thank John Jansen in the Southern Company, Eric Edgerton in the Atmospheric Research & Analysis, Inc, and

researchers involved in ASACA, EPA-STN, IMPROVE, and SEARCH projects for their assistance and hard work, developing the data used.

References

Ames, R. B., and Malm, W. C. (2001). Chemical species' contributions to the upper extremes of aerosol fine mass. *Atmospheric Environment* **35**, 5193-5204.

Boylan, J. W. (2004). PM model performance metrics, goals, and criteria. *Proceeding of 2004 Models-3 conference, Research Triangle, NC.*

Boylan, J. W., Odman, M. T., Wilkinson, J. G., Russell, A. G., Doty, K. G., Norris, W. B., and McNider, R. T. (2005). Integrated assessment modeling of atmospheric pollutants in the Southern Appalachian Mountains. Part I: Hourly and seasonal ozone. *Journal Of The Air & Waste Management Association* **55**, 1019-1030.

Butler, A. J., Andrew, M. S., and Russell, A. G. (2003). Daily sampling of PM2.5 in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. *Journal Of Geophysical Research-Atmospheres* **108**.

Byun, D. W., and Ching, J. K. (1999). Science algorithms of the EPA Models-3 Community Multiscale Air Qualtiy (CMAQ) modeling system. *Atmospheric modeling division, National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711*.

Carter, W. (2000). Documentation of the SAPRC-99 chemical mechanism for VOC reactivity assessment. *Final report to California air resources board contract 92-329 and contract 95-308. Air pollution research center and college of engineering, Center for Environmental Research and Technology, University of California, Riverside, California.*

Emery, C., Tai, E., and Yarwood, G. (2001). Enhanced meteorological modeling and performance evaluation for two texas episodes. *Report to the Texas Natural Resources Conservation Commission, prepared by ENVIRON, International Corp, Novato, CA.*

Gray, H. A., and Cass, G. R. (1998). Source contributions to atmospheric fine carbon particle concentrations. *Atmospheric Environment* **32**, 3805-3825.

Hansen, D. A., Edgerton, E. S., Hartsell, B. E., Jansen, J. J., Kandasamy, N., Hidy, G. M., and Blanchard, C. L. (2003). The southeastern aerosol research and characterization study: Part 1-overview. *Journal Of The Air & Waste Management Association* **53**, 1460-1471.

Hu, Y., Cohan, D. S., Odman, M. T., and Russell, A. G. (2003). Air quality modeling of August 11-20, 2000 for the Fall Line Air Quality Study (FAQS). *Prepared for: Georgia Department of Natural Resources, Environmental Protection Division - Air Protection Branch, Atlanta, GA.*

Jang, C., Phillips, S., Possiel, N., Dolwick, P., B., T., Braverman, T., B., W., M., H., and Fox, T. (2004). Annual model simulations and evaluations of particulate matter and ozone over the continental United States using the Models-3/CMAQ System. *Proceedings of the Models-3 User's workshop, Research Triangle Park. North Carolina*.

Jarvis, C. H., and Stuart, N. (2001). A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part II: The interaction between number of guiding variables and the type of interpolation method. *Journal Of Applied Meteorology* **40**, 1075-1084.

Koch, S. E., Desjardins, M., and Kocin, P. J. (1983). An Interactive Barnes Objective Map Analysis Scheme For Use With Satellite And Conventional Data. *Journal Of Climate And Applied Meteorology* **22**, 1487-1503.

McNair, L. A., Harley, R. A., and Russell, A. G. (1996). Spatial inhomogeneity in pollutant concentrations, and their implications for air quality model evaluation. *Atmospheric Environment* **30**, 4291-4301.

NCAR (2003). PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3. *Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research.*

Nenes, A. (1996). ISORROPIA v1.5: Reference manual. *http://nenes.eas.gatech.edu/ISORROPIA/Version1_5/ISORROPIA15Manual.pdf.*

Nenes, A., Pandis, S. N., and Pilinis, C. (1998). ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols. *Aquatic Geochemistry* **4**, 123-152.

Pace, T. G. (2003). A conceptual model to adjust fugitive dust emissions to account for near source article removal in grid model applications. *US EPA. http://www.epa.gov/ttn/chief/emch/invent/statusfugdustemissions_082203.pdf.*

Park, S.-K., Marmur, A., Ke, L., Yan, B., Russell, A. G., and Zheng, M. (submitted for publication). Source apportionment of PM2.5: Inter-comparison between receptor and emission-based models. *Environmental Science and Technology.*

Park, S.-K., and Russell, A. G. (2003). Sensitivity of PM 2.5 to emissions in the Southeast. *Proceedings of the Models-3 User's workshop, Research Triangle Park, North Carolina.*

Placet, M., Mann, C. O., Gilbert, R. O., and Niefer, M. J. (2000). Emissions of ozone precursors from stationary sources: a critical review. *Atmospheric Environment* **34**, 2183-2204.

Roberts, D. L., and Friedlander, S. K. (1980). Sulfur-Dioxide Transport Through Aqueous-Solutions.2. Experimental Results And Comparison With Theory. *Aiche Journal* **26**, 602-610.

Russell, A. G., McRae, G. J., and Cass, G. R. (1983). Mathematical-Modeling Of The Formation And Transport Of Ammonium-Nitrate Aerosol. *Atmospheric Environment* **17**, 949-964.

Spicer, C. W., Koetz, J. R., Keigley, G. W., Sverdrup, G. M., and Ward, G. F. (1981). A study of nitrogen oxides reactions within urban plumes transported over the ocean. *US-EPA, Research Triangle Park, NC.*

Tilmes, S., and Zimmermann, J. (1998). Investigation on the spatial scales of the variability in measured near-ground ozone mixing ratios. *Geophysical Research Letters* **25**, 3827-3830.

UCAR (2003a). GCIP NCEP Eta model output. *http://dss.ucar.edu/datasets/ds609.2/*.

UCAR (2003b). NCEP ADP Global Surface Obs. *http://dss.ucar.edu/datasets/ds464.0/*.

UCAR (2003c). NCEP ADP Global Upper Air Obs. *http://dss.ucar.edu/datasets/ds353.1/*.

UCAR (2003d). NCEP ADP Global Upper Air Obs. *http://dss.ucar.edu/datasets/ds353.4/*.

UCAR (2003e). TDL U.S. and Canada Surface Hourly Observations. *http://dss.ucar.edu/datasets/ds472.0/*.

Unal, A., Tian, D., Hu, Y., and Russell, A. G. (2003). 2000 Emissions Inventory for Fall Line Air Quality Study (FAQS). *Prepared for: Georgia Department of Natural Resources, Environmental Protection Division - Air Protection Branch, Atlanta, GA.*

US-EPA (2003a). 1999 National Emission Inventory (NEI) documentation and data. *http://www.epa.gov/tth/chief/net/1999inventory.html.*

US-EPA (2003b). Clean air markets -- Data and maps. *http://cfpub.epa.gov/gdm/*.

US-EPA (2003c). Technology transfer network, clearinghouse for inventories and emission factors. *http://www.epa.gov/ttn/chief/emch/spatial/*.

US-EPA (2004a). Biogenic emissions inventory system. *http://www.epa.gov/asmdnerl/biogen.html, ftp://ftp.epa.gov/amd/asmd/beld3, ftp://ftp.epa.gov/amd/asmd/beld3/ascii*.

US-EPA (2004b). Economic growth analysis system version 4.0 and documentation. *http://www.epa.gov/ttn/chief/emch/projection/egas40/*.

US-EPA (2004c). SMOKE User's manual. *http://cf.unc.edu/cep/empd/products/smoke/*.

Watson, J. G., Chow, J. C., Dubois, D., Green, M., Frank, N., and Pitchford, M. (1997). Guidance for network design and optimum site exposure for PM2.5 and PM10. *U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC.*

Zheng, M., Cass, G. R., Schauer, J. J., and Edgerton, E. S. (2002). Source apportionment of PM2.5 in the southeastern United States using solvent-extractable organic compounds as tracers. *Environmental Science & Technology* **36**, 2361-2371.

CHAPTER 4

EVALUATION OF FINE PARTICLE NUMBER CONCENTRATIONS IN CMAQ

(S.-K. Park, A. Marmur, S.B. Kim, D. Tian, P. McMurry, and A. G. Russell.

Aerosol Science and Technology, submitted)

Abstract

The Community Multiscale Air Quality (CMAQ) model is being widely used in air quality management and scientific investigation. Numerous studies have been done for investigating how well CMAQ simulates fine particle mass concentrations but relatively few studies have addressed on how well CMAQ simulated fine particle number distribution. Accurate simulation of particle number concentrations is required because particle number and surface area concentrations may be directly related to human health and visibility. Simulated fine particle number concentrations derived using CMAQ are compared to measurements to identify problems and to improve model performance. Evaluation is done using measured particle number concentrations in Atlanta, Georgia, from 1/1/1999 to 8/31/2000. Assumed particle size of the primary emission in CMAQ cause number concentrations to be significantly underestimated, while particle density has a small impact. Simulating particle size distributions by three lognormal modes cannot accurately simulate particles with size less than 0.01 μm, particularly during nucleation events. An additional mode that accounts for particles smaller than 0.01 μm should be included to improve the accuracy of the number concentration simulations.

Though, the use of the Expectation-Maximization (EM) algorithm to estimate size distribution parameters of measured particles suggests that assumed parameters of lognormal modes in CMAQ are generally reasonable.

4.1. Introduction

Elevated levels of ambient particles have a direct influence on human health; for example, they can trigger reactions ranging from coughing and wheezing to heart attacks and premature death (Donaldson et al., 2001). Particles are also responsible for visibility reduction (Adrian and Orloff, 1977; Faxvog and Roessler, 1978), acid deposition, and perturbing the Earth's radiation balance. In 1997, the National Ambient Air Quality Standards (NAAQS) for fine particle mass concentrations were promulgated in response to a growing number of scientific linkages between fine particle levels and adverse health effects. In addition to studies associating particle mass concentrations with the human health (Brauer et al., 2001; Dockery and Pope, 1994; Dockery et al., 1993; Ebelt et al., 2000; Gold et al., 1999; Pope et al., 1995; Pope et al., 1999; Saldiva et al., 1995; Vedal, 1997), some studies suggest that fine particle number and surface area concentrations are more directly related with pulmonary disease than particle mass concentrations (Donaldson et al., 1998; Penttinen et al., 2001; Woo et al., 2001b). In addition, particle mass concentrations are not always correlated with particle number concentrations, for example, when very small particles are formed by nucleation, particle mass does not significantly increase (Woo et al., 2001a). Hence, if particle number or surface area concentrations are responsible for human health, strategies designed to meet NAAQS for $PM_{2.5}$ might not be as effective for protecting human health as desired (Woo et al.,

2001a). These studies also suggested that the particle number concentration itself was an important index in epidemiological analyses.

The physics, chemistry, and the response of particles to emissions have been studied using air quality models. One of the most comprehensive air quality modeling frameworks is Models-3, composed of the Community Multiscale Air Quality (CMAQ) model (Byun and Ching, 1999), the Mesoscale meteorological Model version 5 (MM5) (PSU/NCAR, 2003), and Sparse Matrix Operator Kernel Emissions (SMOKE) (US-EPA, 2004e). CMAQ predicts gas-phase species as well as particle concentrations. Particles in CMAQ are divided into two groups, fine ($PM_{2.5}$: particles with size less than 2.5 μm) and coarse particles (PM_{10}: particles with size between 2.5 and 10 μm).

$PM_{2.5}$ is treated by two interacting sub-distributions (or modes). Aitken (smaller one) mode represents fresh particles either from nucleation or from direct emissions, while accumulation (larger one) mode represents aged particles. The production rate of new particle number, J [m^{-3} sec^{-1}], is calculated from the nucleation of sulfuric acid vapor, assuming the new particle diameter is 2.0 nm based the classical binary homogeneous nucleation theory (Wilemski, 1984) incorporating the hydration effect (Kulmala et al., 1998). The two modes interact with each other through coagulation, and particles in Aitken mode may grow into accumulation mode. PM_{10} is treated by a coarse mode originated from the direct emissions by natural processes such as wind blowing dust, or sea salt particles by wind driven waves on the sea surface. Chemical species of $PM_{2.5}$ treated in CMAQ are fine species sulfate, nitrate, ammonium, water, anthropogenic and biogenic organic carbon, elemental carbon, and other unspecified material of

anthropogenic origin. PM_{10} includes sea salt, wind-blown dust, and other unspecified material of anthropogenic origin.

Fine particle mass concentrations simulated using CMAQ have been extensively evaluated using relatively abundant measurement data (Arnold et al., 2003; Eder et al., 2003; Jun and Stein, 2004; Mebust et al., 2003; Park et al., 2005a; Tonnesen, 2003). However, CMAQ's ability to reproduce aerosol number concentrations has not been addressed in such detail. A few studies have found that CMAQ under predicts aerosol number concentrations of the Aitken mode (i.e., particles with size less than 0.1 μm) by a factor of 10 to 100, and those of the accumulation mode (i.e., less than about 2 μm in the diameter) by a factor of 5 to 10 for the Pacific Northwest (Elleman et al., 2004, 2005). Errors in particle number concentrations were significantly larger than those of particle mass concentrations or gas-phase constituents. Thus, the authors concluded that the modeled mean particle size in CMAQ might be significantly larger than the measured mean size due to the inadequate treatment of aerosols as three lognormal modes or the inaccurate process of the aerosol size distribution. However, the study did not further analyze the sources of the errors.

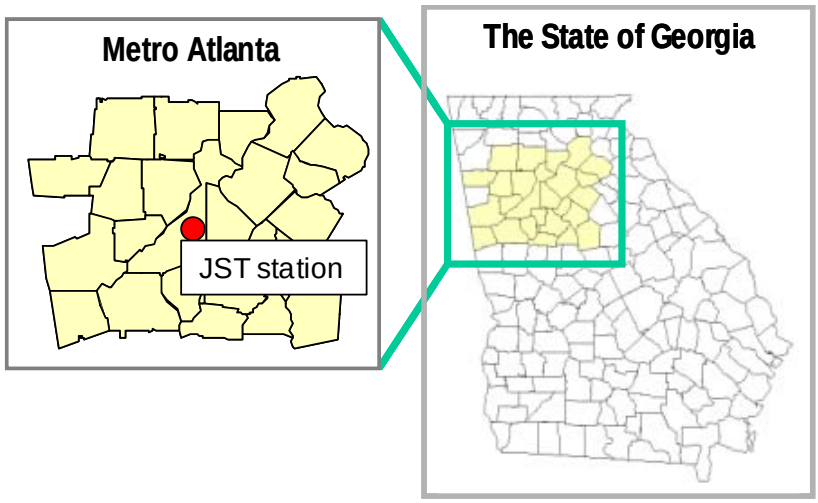
The goal of this research is to evaluate fine particle number concentrations simulated by CMAQ. The evaluation uses hourly measured particle number concentrations of 39 size bins from 3 nm to 3 μm for one year and eight months from 1/1/1999 to 8/31/2000 at an urban area located in Atlanta, Georgia. In addition, this paper extensively analyzes the causes of errors, and examines the feasibility to improve the accuracy of the simulations.

4.2. Methods

Fine particle number and mass levels were monitored at the Jefferson Street (JST) monitor in urban Atlanta, Georgia (Figure 4-1a) from 1/1 1999 to 8/31/2000 as a part of the Southeastern Aerosol Research and Characterization (SEARCH) study (Woo et al., 2001a). Hourly particle number concentrations were measured for 39 size bins from 3 nm to 3 μm in diameter using a nano-scanning mobility particle sizer (NSMPS) (3 nm – 50 nm), a standard scanning mobility particle sizer (SMPS) (20 nm – 250 nm), and a laser particle counter (LPC) (100 nm – 2 μm) (Van Loy et al., 2000; Woo et al., 2001a). Hourly $PM_{2.5}$ mass concentrations were measured using the Tapered Element Oscillating Microbalance (TEOM) (Patashnick and Rupprecht, 1991).

Hourly particle number and mass concentrations were simulated using EPA's Models-3 framework: SMOKE (v1.5), MM5 (v3.5.3), and CMAQ (v4.3). CMAQ was run from 1/1/1999 to 8/31/2000 over the eastern United States with 78*66 grids of size 36 km (Figure 4-1b) (Marmur et al., 2004). Detailed information of the parameters used in the modeling system is available in elsewhere (Park et al., 2005a). Meteorological fields were evaluated using the TDL surface hourly data (UCAR, 2003e), which were not used as an input of MM5. Mean errors (MEs) in temperature, specific humidity, and wind speed were 1.5 ^{o}C, 1.8g kg^{-1}, and 0.8m sec^{-1}, respectively. These values are within the benchmarks for the metrological model evaluation (Emery et al., 2001).

(a)



(b)

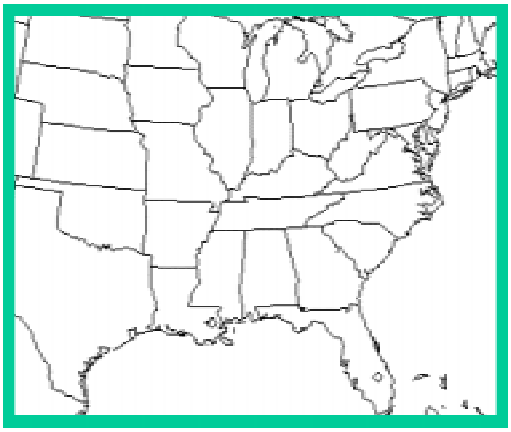


Figure 4-1. (a) Particle number and mass monitor (JST station) in Atlanta, Georgia. (b) Air quality model (CMAQ) domain.

Particle number concentrations are measured in each of 39 size bins, but CMAQ concentrations are simulated using three modes, which make a direct comparison cumbersome. The three modes, Aitken(i), accumulation(j), and coarse(c), follow a lognormal distribution. To compare the simulated concentrations with the measured concentrations, modal particle number concentrations from CMAQ were converted into size resolved particle number concentrations expressed in the aerodynamic diameter using PMx (Jiang, 2005; Jiang et al., 2004; Jiang and Yin, 2001). PMx estimates the parameters of the lognormal distribution in each mode using particle number, surface area, and mass concentrations. PMx calculates particle number concentrations in each size bin by the integral of the probability density function of particle number concentrations for each interval. Accuracy of the particle concentrations in each size bin is dependent on the simulated particle number, surface area, and mass concentrations.

Parameters of the lognormal distributions can be also obtained from CMAQ's diagnostic output variables (Bhave and Binkowski, 2005). Parameters of the lognormal distributions differ (Table 4-1), and the two methods calculate the size resolved particle number concentrations differently. Number concentrations for particles with size between 3 nm to 0.5 μm are a little higher when PMx is used, and concentrations for particles with size between 0.5 μm to 2.5 μm are somewhat higher when CMAQ's diagnostic output variables are used. However, the difference is negligible for total particle number concentrations with size less than 2.5 μm, and those differences do not significantly affect the results of the analysis performed in this paper. Also, size resolved particle number concentrations are calculated from modal concentrations from CMAQ based on the aerodynamic diameter. Size resolved concentrations differ slightly when using other

kinds of diameters, such as Stokes or electric mobility diameter. Differences would not significantly influence on analysis.

Table 4-1. Geometric mean (D_g) and geometric standard deviation (σ_g) of lognormal distributions obtained from PMx and CMAQ's diagnostic variables from 3/23/1999 to 3/30/1999 at the JST station in Atlanta, Georgia.

		Mode		
		Aitken	Accumulation	Coarse
PMx	D_g [μm]	0.08	0.11	0.81
	σ_g	1.83	2.07	2.20
Diagnostic output	D_g [μm]	0.05	0.07	0.80
	σ_g	2.52	2.66	2.20

4.3. Results

Simulated fine particle number concentrations were significantly lower than measured concentrations by a factor of up to 1000 (Figure 4-2). Note that the y-axis is in a logarithmic scale. Simulated surface area concentrations were also much lower than measured values, whereas simulated volume and mass concentrations reasonably well match the measured values. In addition, CMAQ underestimated number concentrations for size less than 1 μm, and overestimated concentrations for size larger than 1 μm (Figure 4-3). These results suggested that CMAQ overestimated particle sizes.

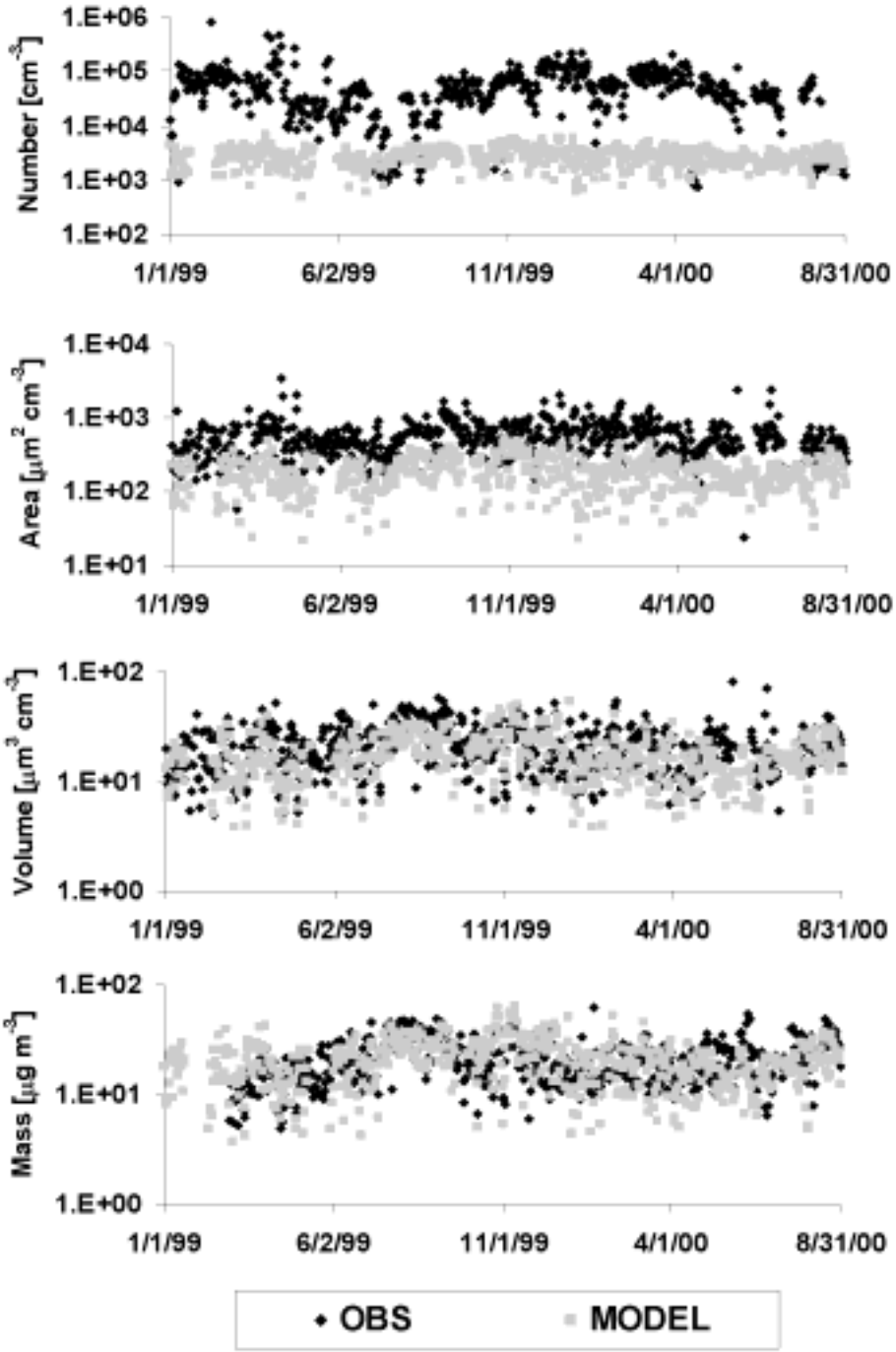


Figure 4-2. Daily particle number, surface area, volume, and mass concentrations for particles less than 2.5 μm from 1/1/1999 to 8/31/2000 at the JST station in Atlanta, Georgia.

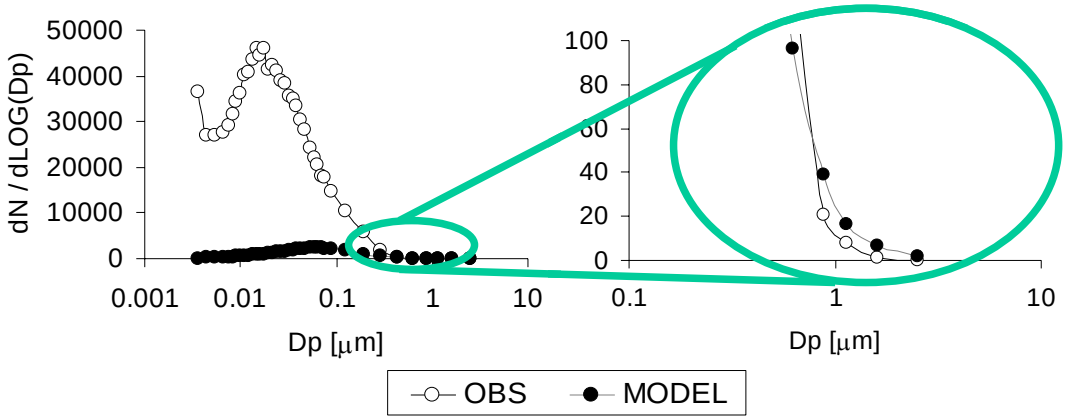


Figure 4-3. Median number concentrations of particles for 39 size bins (from 3 nm to 2.5 μm) from 1/1/1999 to 8/31/2000 at the JST station in Atlanta, Georgia.

The discrepancy between measured and simulated particle number concentrations could be partly due to the different spatial scales between the observation and the model (Park et al., 2005a). Measured concentrations are point values at the monitoring station, whereas CMAQ simulates volume-averaged concentrations. To analyze the effect of the different spatial scales, particle number concentrations were simulated using 12 km grids as well as 36 km grids from July 1-31, 2001 in a separate study (Park and Russell, 2003a). Number concentrations at the JST station from the 12 km grid were 17% higher than those from the 36 km. Thus, a small part of the errors in simulated number concentrations could be from the different spatial scales.

Temporal trends of the simulated number concentrations were also compared with that of measured concentrations. Measured particle concentrations with size less than 45 nm in the diameter were relatively higher for winter and spring, but simulated

concentrations did not have an apparent seasonal trend (Figure 4-4 (a)). Also, measured concentrations had a marked weekly variation, which is not strong in the simulated concentrations (Figure 4-4 (b)). In addition, the diurnal peak coinciding with the morning and afternoon rush hours for particles from 1 μm to 2 μm were observed in measured concentrations, but not in those simulated (Figure 4-4 (c)). Thus, measured number concentrations had more apparent seasonal/daily/diurnal variations than simulated concentrations. The relatively smooth temporal variation of simulated number concentrations was consistent with that of simulated mass concentrations, which is caused in part by the smooth emission profiles used in CMAQ, the spatial averaging effect, and the spatial allocation of emissions (Marmur et al., 2005a; Park et al., 2005c; Park et al., 2005d) as well as the inability of the meteorological model to simulate the very fine features (Hogrefe et al., 2004). In summary, analysis of simulated number concentrations for one year and eight months showed that simulated number concentrations were significantly underestimated, and had less temporal variation.

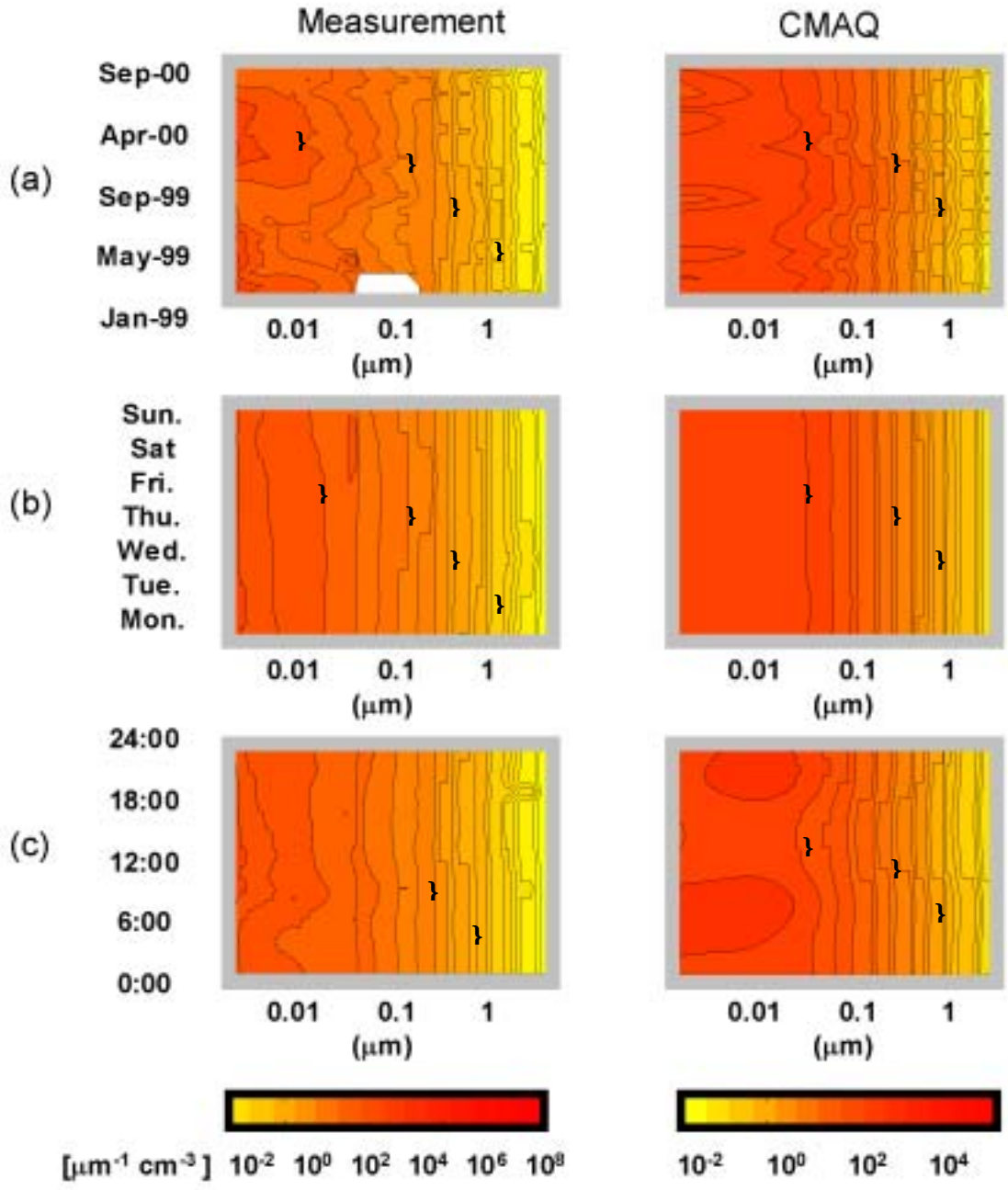


Figure 4-4. Temporal trends of the particle number concentrations from 1/1/1999 to 8/31/2000 at the JST station in Atlanta, Georgia. (a) monthly, (b) day of week, and (c) hourly average concentrations.

4.4. Analysis

Underestimation of particle number and surface area is due in part to the assumptions used. Here, three assumptions: particle density, size distribution of primary emissions, and the treatment of aerosols as three lognormal modes, are examined.

4.4.1 Particle Density

CMAQ converts particle mass concentrations into particle number concentrations using assumed particle density information (Table 4-2) (Byun and Ching, 1999). Assumed particle densities of organic carbon and elemental carbon in CMAQ appear larger than measured values. Hence, simulated number concentrations in CMAQ would be somewhat underestimated if simulated mass concentrations were reasonably well matched with the measured values, and the size distributions based on particle mass are correct as well.

Table 4-2. Assumed and measured particle densities (g cm^{-3}).

	CMAQ	Literature
sulfate, nitrate, and ammonium	1.8	1.77 [a], 1,72 ~ 1.83 [d]
elemental carbon	2.2	1.77 [b], 2.0 [d]
organic carbon	2	1.2 ~ 1.78 [c], 1.2 [d]

a (Zhang et al., 2005)
b (Park et al., 2004a; Park et al., 2004b)
c (Turpin and Lim, 2001)
d (McMurry et al., 2002)

Influence of the inaccurate density on estimating particle number concentrations was examined by comparing the difference between measurements and simulations. Average ratios of measured to simulated number concentrations were 1.15 (mass), 1.43 (volume), 3.95 (surface area), and 29.34 (number) (Table 4-3). Note that $PM_{2.5}$ mass measured by TEOM contains small amounts of water, although some studies suggest that the amount of water would be negligible (Eatough et al., 2003; Gong and Demerjian, 1995; Meyer et al., 2000; Schwab et al., 2004). Considering the possibility of water in the particle, the true ratio of observed mass to modeled mass would be slightly less than 1.15. The ratio increased from mass to volume concentrations, which is due to the higher particle densities (especially, OC and EC) used in CMAQ (see Table 4-2).

Table 4-3. Summary statistics of hourly average particle number, surface area, volume, and mass concentrations for particles less than 2.5 μm from 1/1/1999 to 8/31/2000 at the JST station in Atlanta, Georgia

	Number [cm^{-3}]		Area [$\mu m^2\ cm^{-3}$]		Volume [$\mu m^3\ cm^{-3}$]		Mass [$\mu g\ m^{-3}$] (water was excluded)	
	OBS	MODEL	OBS	MODEL	OBS	MODEL	OBS	MODEL
Mean	7.5E+04	2.5E+03	663.7	168.1	22.4	15.7	21.3	18.5
STDEV	9.5E+04	1.3E+03	549.1	101.7	13.4	9.6	13.0	10.6
Minimum	2.2E+03	7.6E+01	48.4	0.9	1.5	0.1	0.9	0.0
25th percen.	3.2E+04	1.6E+03	369.9	94.7	12.8	8.9	11.9	10.9
Median	5.2E+04	2.3E+03	529.1	151.0	18.9	14.0	18.3	17.0
75th percen.	8.8E+04	3.3E+03	781.7	222.2	28.6	20.6	27.8	24.4
Maximum	1.8E+06	1.1E+04	8533.7	827.9	127.7	79.5	97.5	81.4

Particle densities in CMAQ are changed to the measured particle densities, and particle mass and number concentrations were calculated with CMAQ (Table 4-2). When densities of OC and EC in CMAQ are assumed as 1.2 g cm^{-3} and 2.0 g cm^{-3}, respectively (McMurry et al., 2002), the number concentrations increased 10.7%. When densities of OC and EC are assumed as 1.78 and 1.77, respectively (Park et al., 2004a; Park et al., 2004b; Turpin and Lim, 2001), the number concentrations increased 10.9%. However, mass concentrations changed less than 1% in both cases. Thus, CMAQ under predicted particle number concentrations 10% due to the large particle density, but the particle density alone cannot fully explain the large difference between measured and simulated number concentrations.

4.4.2. Size Distribution of Primary PM Emissions

Measured and simulated particle concentrations for different particle sizes also can be compared. CMAQ overestimated particle number concentrations for large particles (larger than 1.0 μm), but significantly underestimated particle number concentrations for small particles (less than 1.0 μm) (Figures 4-3 and 4-4). This result was expected because simulated number concentrations were significantly lower than the measured although simulated volume concentrations were predicted reasonably well (Figure 4-2). To investigate why this might be, the influence of the size distribution of initial emissions on particle size distributions was analyzed.

The EPA emission inventory for particulate matter does not contain information about source size distributions. In CMAQ, the default assumption is that the major fraction of $PM_{2.5}$ emissions are in the accumulation mode with a small fraction in the

Aitken mode; i.e., 99.9% of the fine primary organic aerosol (POA) and primary elemental carbon (PEC) emissions are assumed to be in the accumulation mode, and the remaining fraction, 0.1%, are assigned to the Aitken mode (Byun and Ching, 1999), and 100% of primary sulfate (PSO4), primary nitrate (PNO3), and unspecified fine PM (PMFINE) emissions are assigned to the accumulation mode. However, studies show that around 85% of primary organic carbon (POA) and elemental carbon (PEC) emissions should be assigned to the accumulation mode, and the remaining 15% of POA and PEC to the Aitken mode (Venkataraman et al., 1994). If so, the assumption that 99.9% of POA and PEC emissions would be in the accumulation mode needs modification. Currently, other air quality models, for example, the European Monitoring and Evaluation Program (EMEP) unified aerosol model version, assume that 85% of POA and PEC are in the accumulation mode, and 100% of PMFINE is in the accumulation mode (Tsyro, 2002).

Sensitivity analysis of how simulated particle number concentrations respond to the ratio of the primary PM emissions in the accumulation to Aitken modes for $PM_{2.5}$ emissions was performed for a limited period (3/23/1999 to 3/30/1999), which includes both regular (3/23 ~ 3/28) and nucleation (3/29 ~ 3/30) days, the latter being characterized by exceptionally high concentrations of particles with size less than 45 nm due to nucleation (Figure 4-5) (Woo et al., 2001a). The model was run with the ratio of 0.85/0.15 (accumulation/Aitken), 0.9/0.1, and 0.99/0.01, in addition to the default ratio of 0.999/0.001 for PEC and POA emissions. When the ratio changed from 0.999/0.001 to 0.85/0.15, simulated particle number concentrations increased by a factor of 6 (Table 4-4), but simulated particle mass concentrations remain virtually the same (less than 1% change). Thus, as the ratio changes from 0.999/0.001 to 0.85/0.15, the simulated number

concentrations approached measured concentrations for the regular days, but number concentrations were still lower than those for the nucleation days.

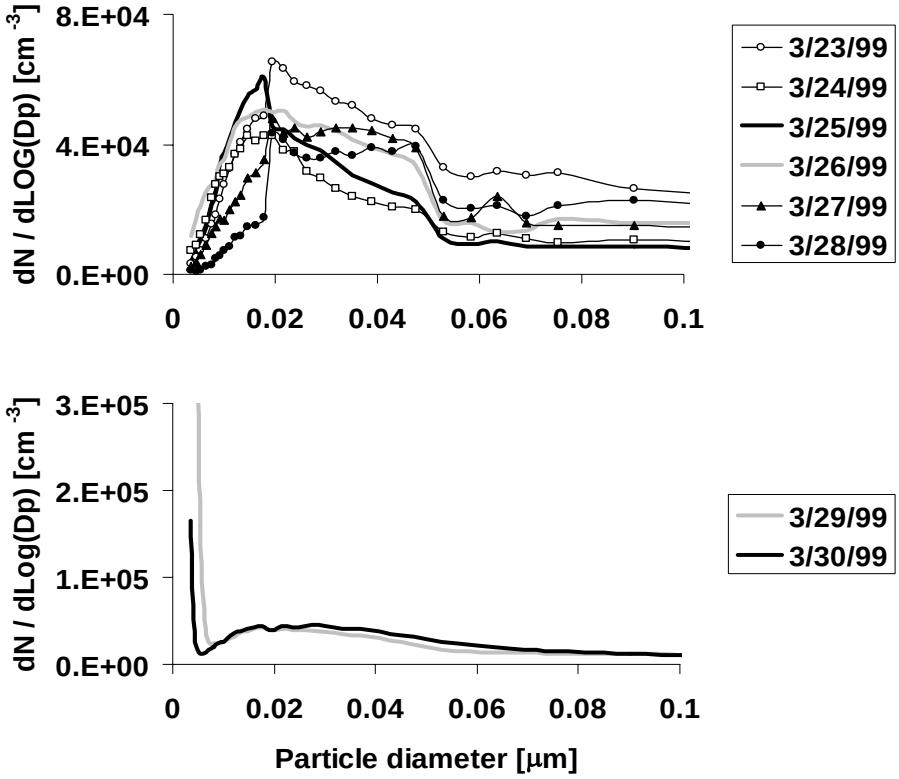


Figure 4-5. Daily average measured particle number concentrations from 3/23/1999 to 3/30/1999 at the JST station in Atlanta, Georgia.

Table 4-4. Simulated fine particle number concentrations by CMAQ, as initial ratio of particle emissions in Aitken (E_i) and accumulation (E_j) modes change from 3/23/1999 to 3/30/1999 at the JST station in Atlanta, Georgia.

	$E_i : E_j$	Number [cm^{-3}]
CMAQ	0.001:0.999 (default)	4.2E+03
	0.01:0.99	5.9E+03
	0.10:0.90	1.8E+04
	0.15:0.85	2.5E+04
OBS	Regular days	4.5E+04
	Nucleation event days	1.8E+05

In summary, this analysis showed that the assumed size distribution of primary particle emissions significantly affects the number concentrations, but the assumed size distribution itself did not completely explain the underestimation of simulated particle number concentrations by CMAQ. The size distribution of particle emissions is a hot topic of the current issue (McMurry and Woo, 2002), and still further research is needed. Also, results analyzed in this section show that number concentrations predicted by CMAQ are much lower than measured values for nucleation events days even when size distribution of primary emissions are changed. Studies show that nucleation that occurs when hot combustion emissions mix with cool ambient air can substantially increase number concentrations associated with vehicle exhaust, and the tendency of vehicle exhaust to nucleate increases as ambient temperatures decrease (Tobias et al., 2001). However, the exact physics and chemistry leading to such phenomena is still uncertain (Gaydos et al., 2005; Kulmala et al., 2004).

4.4.3. Treatment of Aerosol as Three Lognormal Modes

CMAQ simulates particles assuming that the particle size distribution follows three lognormal distributions: Aitken, accumulation, and coarse modes (Byun and Ching, 1999; Whitby, 1978a). Distribution parameters in CMAQ are summarized in Table 4-5. To evaluate the distribution parameters used, measured data were fit to lognormal distributions (Figure 4-6) using an Expectation-Maximization (EM) algorithm. The EM algorithm is described in the Appendix to estimate these parameters. Coarse mode parameters could not be evaluated because size measurements were not made above 3 μm.

Table 4-5. Parameters of the lognormal distributions for the particle number, surface area, and volume distributions in CMAQ, and those in the measurements based on the median concentrations from 1/1/1999 to 8/31/2000 at the JST station in Atlanta, Georgia.

		CMAQ*		OBS**	
variable	mode	geometric mean (μm)	geometric standard deviation	geometric mean (μm)	geometric standard deviation
Number (n)	Aitken	0.01	1.7	0.017	2.34
	accumulation	0.07	2.0	0.07	1.88
	coarse	1.00	2.2	NA	NA
Surface area (s)	Aitken	0.02	1.7	0.043	2.03
	accumulation	0.18	2.0	0.18	1.81
	coarse	3.47	2.2	NA	NA
Volume (mass) (v)	Aitken	0.02	1.7	0.045	2.03
	accumulation	0.30	2.0	0.28	2.03
	coarse	6.46	2.2	NA	NA

* Parameters for the surface area and mass (volume) concentrations are calculated as follows
  · Surface area: $D_{gs}=\ln D_{gn} + 2 (\ln \sigma_{gn})^2$, $\sigma_{gs}=\sigma_{gn}$
  · Volume (mass): $D_{gv}=\ln D_{gn} + 3 (\ln \sigma_{gn})^2$, $\sigma_{gs}=\sigma_{gn}$
** Refer Figure 6

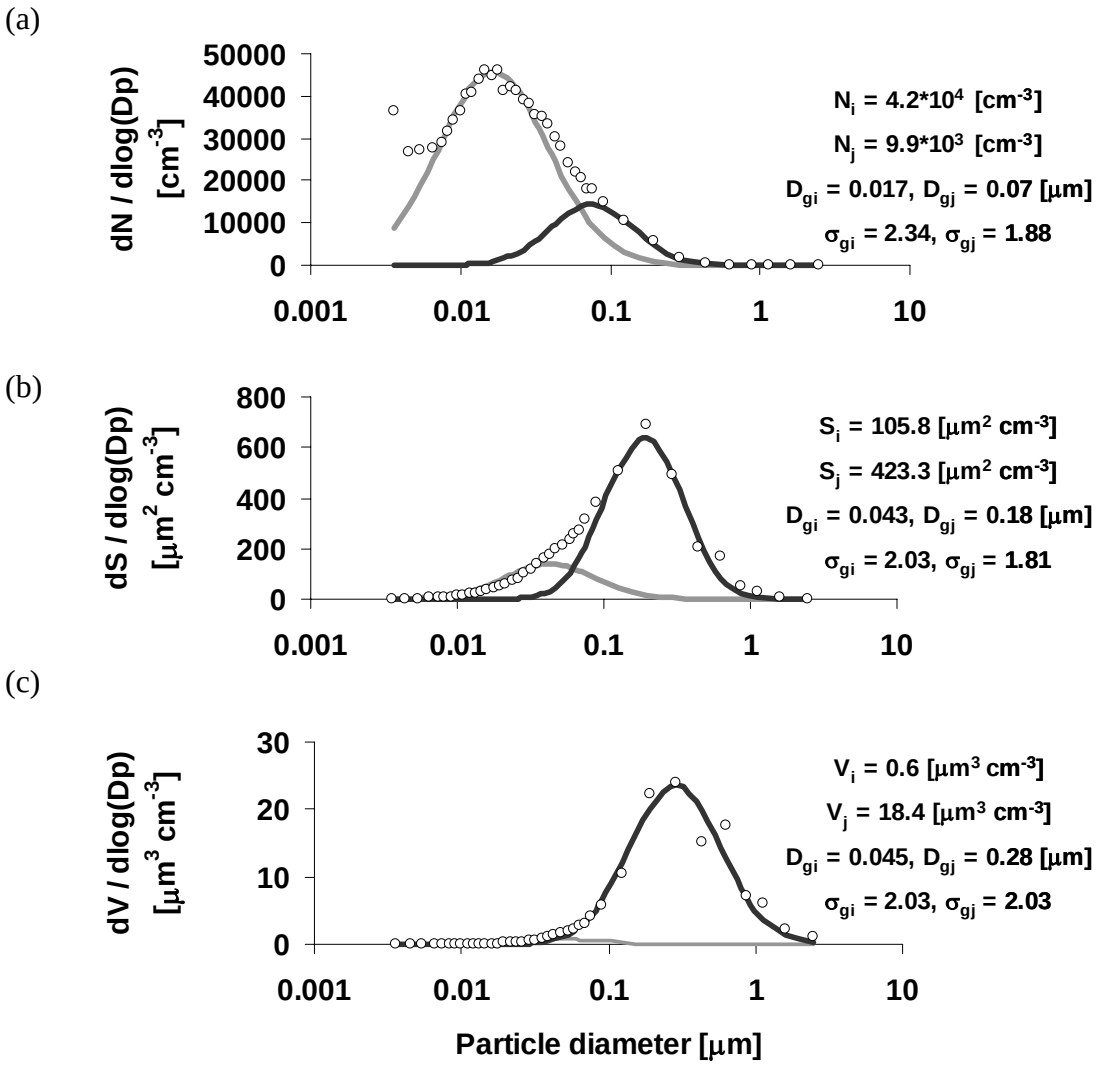


Figure 4-6. (a) Measured median particle number, (b) surface area, and (c) volume concentrations from 1/1/1999 to 8/31/1999 at the JST station in Atlanta, Georgia, and the fitted distributions using parameters from the EM algorithm along with the number (N), surface area (S), and volume (V) concentrations, and the geometric mean (D_g), and geometric standard deviation (σ_g) for Aitken (i), and accumulation (j) modes.

The distribution using the lognormal parameters from the EM algorithm fit well the measured number concentrations for particles larger than 0.01 μm, but the measured number concentrations did not fit well for particles with size less than 0.01 μm (Figure 4-6a). In order to better predict particles with size less than 0.01 μm, an additional mode with the geometric mean diameter of which is smaller than that of the Aitken mode, i.e., a nucleation mode should be added in CMAQ. Accurate prediction of nucleated particles is also important in predicting particles in Aitken mode because nucleated particles grow into Aitken mode. Thus, parameters estimated using the measured data are slightly different with those in CMAQ for Aitken mode partly due to the inaccurate simulation of nucleated particles, although those in accumulation mode are similar with each other (Table 4-5). In summary, the treatment of aerosols as three lognormal modes should simulate particles larger than 0.01 μm reasonably well, but does not accurately simulate particles smaller than 0.01 μm.

Currently, the EMEP aerosol model version includes a nucleation mode, in addition to the Aitken, accumulation, and coarse modes (Tsyro, 2002). Performance of predicting particle number concentrations of the multi (four) distribution model was significantly better than that of the mono-disperse model (Tsyro, 2002). Studies suggest that the absolute rate of the particle formation used in CMAQ is uncertain, so using relative rates of new particle formation and condensation on existing particles is less uncertain than the absolute rates used in the current version (McMurry, 1983; McMurry and Friedlander, 1979). Currently, a new aerosol module, the Model of Aerosol Dynamics, Reaction, Ionization, and Dissolution (MADRID), calculates the relative rates of new particle formation and condensation on existing particles (Zhang et al., 2004).

4.5. Conclusions

Fine particle mass concentrations simulated using CMAQ have been extensively evaluated, but CMAQ's ability to reproduce aerosol number concentrations is not as well demonstrated. Particle number concentrations simulated by CMAQ from 1/1/1999 to 8/31/2000 were evaluated. Analysis showed that the predicted number concentrations were significantly underestimated by a factor of up to 1000. Influences of assumptions used in CMAQ on the predicted particle number concentrations were investigated, including impacts of particle density, size distribution of the initial emissions, and the treatment of aerosols as three lognormal modes.

One cause of underestimating particle number concentrations was that the assumed particle density in CMAQ is higher than the observed values though this had a small impact. Another cause of the underestimation was the assumed size distribution of initial particle emissions in CMAQ was larger than the observed. Currently, CMAQ assigns only 0.1% of fine particle emissions to the Aitken mode, and the remaining 99.9% of fine particle emissions to the accumulation mode. However, measurements show that around 15% of POA and PEC emissions should be treated as being in the Aitken mode, and the remaining 85% as the accumulation mode. Sensitivity test results showed that when the fraction of initial particle emissions in the Aitken mode increased, particle number concentrations increased and approached measured values. Another important cause of the underestimation was the treatment of aerosols using three lognormal distributions. The size distribution of the measured concentrations showed that currently used three lognormal distributions could not accurately capture particles with size less than 0.01 μm. An additional mode that accounts for the nucleated particles

should be included to improve the accuracy of the number concentration simulations. Further work on parameterizing nucleation is also needed.

Appendix

An Expectation-Maximization (EM) algorithm was employed to estimate the parameters of a mixture of lognormal distributions. The EM algorithm is based on the concept of incomplete data for optimizing likelihood functions (Demster et al., 1977; Redner and Walker, 1984). When estimating the parameters of mixture distributions, we posit that the observed data *X* are incomplete and consider the existence of unobserved data, *Y*. This significantly facilitates optimizing the likelihood function since the optimizing process is analytically intractable in many cases without assuming *Y*. More precisely, let's assume the following mixture lognormal distributions.

$$f(X \mid \Phi) = \sum_{i=1}^{K} \alpha_i \, p_i(X \mid \theta_i),$$

where $p_i(X \mid \theta_i) = \frac{1}{\sqrt{2\pi}\sigma_i X} e^{-\frac{[\ln(X)-\mu_i]^2}{2\sigma_i^2}}$,

and $\Phi = \{\alpha_i, \theta_i\}$, $\theta_i = \{\mu_i, \sigma_i^2\}$ for $i=1, \ldots, K$ such that $\sum_{i=1}^{K} \alpha_i = 1$.

Parameters of mixture lognormal distributions include the number of components in the mixtures, their coefficients, means, and variances denoted *K*, α_i μ_i, and σ_i^2, respectively. *K* is usually obtained by prior knowledge. Here, we set *K*=2 as the number of particles that are typically divided into two modes for size less than 2.5 μm (Whitby, 1978b). The EM algorithm is performed by repeating between the so-called the E-step

and the M-step. In the E-step, we compute the expectation of complete log-likelihood function conditional on the initial guess of parameters and the observed data called the Q function.

$$Q(\Phi, \Phi^{(i-1)}) = E[\log p(X, Y \mid \Phi) \mid X, \Phi^{(i-1)}]$$

In the M-step, we update the estimate of parameters to maximize the function computed in E-step.

$$\Phi^{(i)} = \arg\max_{\Phi} Q(\Phi, \Phi^{(i-1)})$$

We then obtain a set of updated parameters by iterating between the E-step and the M-step. Final parameters are determined when the difference between $\Phi^{(i)}$ and $\Phi^{(i+1)}$ is less than the specified threshold or satisfied with the specified number of iterations. More details of the EM algorithm is available elsewhere (Hogg et al., 2005).

Acknowledgement

This research was supported by the U.S. Environmental Protection Agency under Agreements RD82897602, RD83107601, and RD83096001.

References

Adrian, R. J., and Orloff, K. L. (1977). Laser Anemometer Signals - Visibility Characteristics And Application To Particle Sizing. *Applied Optics* 16, 677-684.

Arnold, J. R., Dennis, R. L., and Tonnesen, G. S. (2003). Diagnostic evaluation of numerical air quality models with specialized ambient observations: testing the Community Multiscale Air Quality modeling system (CMAQ) at selected SOS 95 ground sites. *Atmospheric Environment* 37, 1185-1198.

Bhave, P., and Binkowski, F. S. (2005). Private communication.

Brauer, M., Ebelt, S. T., Fisher, T. V., Brumm, J., Petkau, A. J., and Vedal, S. (2001). Exposure of chronic obstructive pulmonary disease patients to particles: Respiratory and cardiovascular health effects. *Journal Of Exposure Analysis And Environmental Epidemiology* 11, 490-500.

Byun, D. W., and Ching, J. K. (1999). Science algorithms of the EPA Models-3 Community Multiscale Air Qualtiy (CMAQ) modeling system. *Atmospheric modeling division, National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711*.

Demster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society (B)* 39, 1-38.

Dockery, D. W., and Pope, C. A. (1994). Acute Respiratory Effects of Particulate Air-Pollution. *Annual Review of Public Health* 15, 107-132.

Dockery, D. W., Pope, C. A., Xu, X. P., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., and Speizer, F. E. (1993). An Association Between Air-Pollution And Mortality In 6 United-States Cities. *New England Journal Of Medicine* 329, 1753-1759.

Donaldson, K., Li, X. Y., and MacNee, W. (1998). Ultrafine (nanometre) particle mediated lung injury. *Journal Of Aerosol Science* 29, 553-560.

Donaldson, K., Stone, V., Seaton, A., and MacNee, W. (2001). Ambient particle inhalation and the cardiovascular system: Potential mechanisms. *Environmental Health Perspectives* 109, 523-527.

Eatough, D. J., Long, R. W., Modey, W. K., and Eatough, N. L. (2003). Semi-volatile secondary organic aerosol in urban atmospheres: meeting a measurement challenge. *Atmospheric Environment* 37, 1277-1292.

Ebelt, S. T., Petkau, A. J., Vedal, S., Fisher, T. V., and Brauer, M. (2000). Exposure of chronic obstructive pulmonary disease patients to particulate matter: Relationships between personal and ambient air concentrations. *Journal Of The Air & Waste Management Association* 50, 1081-1094.

Eder, B., Yu, S., and Dennis, R. (2003). An Evaluation of the Models-3 CMAQ Aerosol Module. *PM 2003 Meeting -- Particulate Matter: Atmospheric Sciences, Exposure and the Forth Colloquium on PM and Human Health, Pittsburgh, PA.*

Elleman, R. A., Kotchenruther, R. A., Covert, D. S., Mass, C. F., and J., C. (2004). CMAQ aerosol number and mass evaluation for Pacific Northwest. *Models-3 User's workshop, Research Triangle Park, North Carolina.*

Elleman, R. A., Kotchenruther, R. A., Covert, D. S., Mass, C. F., and J., C. (2005). CMAQ aerosol number and mass evaluation for Pacific Northwest. *7th conference on atmospheric chemistry, The 85th AMS annual meeting, San Diego, CA.*

Emery, C., Tai, E., and Yarwood, G. (2001). Enhanced meteorological modeling and performance evaluation for two texas episodes. *Report to the Texas Natural Resources Conservation Commission, prepared by ENVIRON, International Corp, Novato, CA.*

Faxvog, F. R., and Roessler, D. M. (1978). Carbon Aerosol Visibility Vs Particle-Size Distribution. *Applied Optics* 17, 2612-2616.

Gaydos, T. M., Stanier, C., and Pands, S. (2005). Modeling of in situ ultrafine atmospheric particle formation in the eastern United States. *Journal Of Geophysical Research-Atmospheres* 110, doi:10.1029/2004JD004683.

Gold, D. R., Damokosh, A. I., Pope, C. A., Dockery, D. W., McDonnell, W. F., Serrano, P., Retama, A., and Castillejos, M. (1999). Particulate and ozone pollutant effects on the respiratory function of children in Southwest Mexico City (vol 10, pg 8, 1999). *Epidemiology* 10, 470-470.

Gong, Q., and Demerjian, K. L. (1995). Hydrocarbon Losses On A Regenerated Nafion(R) Dryer. *Journal Of The Air & Waste Management Association* 45, 490-493.

Hogg, R. V., Craig, A. T., and McKean, J. W. (2005). Introduction to Mathematical Statistics. *Prentice Hall, New York.*

Hogrefe, C., Biswas, J., Lynn, B., Civerolo, K., Ku, J. Y., Rosenthal, J., Rosenzweig, C., Goldberg, R., and Kinney, P. L. (2004). Simulating regional-scale ozone climatology over the eastern United States: model evaluation results. *Atmospheric Environment* 38, 2627-2638.

Jiang, W. (2005). Differences Between CMAQ Fine Mode Particle and PM2.5 Concentrations and their Impact on Model Performance Evaluation in the Lower Fraser Valley. *Atmospheric Environment.*

Jiang, W., Giroux, E., Roth, H., and Yin, D. (2004). Evaluation of CMAQ PM results using size-resolved field measurement data: the particle diameter issue and its impact on model performance assessment. *ITM conference*.

Jiang, W., and Yin, D. (2001). Development and application of the PMx software package for converting CMAQ modal particulate matter results in size-resolved quantities. *Prepared for Program of Energy Research and Development, Natural Resources Canada. Report number: PET-1497-01S, 2001.*

Jun, M., and Stein, M. L. (2004). Statistical comparison of observed and CMAQ modeled daily sulfate levels. *Atmospheric Environment* 38, 4427-4436.

Kulmala, M., Laaksonen, A., and Pirjola, L. (1998). Parameterizations for sulfuric acid/water nucleation rates. *Journal Of Geophysical Research-Atmospheres* 103, 8301-8307.

Kulmala, M., Vehkamaki, H., Petajda, T., Dal Maso, M., Lauri, A., Kerminen, V. M., Birmili, W., and McMurry, P. H. (2004). Formation and growth rates of ultrafine atmospheric particles: a review of observations. *Journal Of Aerosol Science* 35, 143-176.

Marmur, A., Mulholland, J., and Russell, A. (2004). Air quality modeling of PM2.5 mass and composition in Atlanta: Results from a two-year simulation and implications for use in health studies. *Models-3 User's workshop, Research Triangle Park, North Carolina*.

Marmur, A., Park, S.-K., Mulholland, J., and Russell, A. G. (2005). PM2.5 Source Apportionment Using Receptor and Source-oriented Models: Conceptual Differences and Implications for Health Studies. *Proceedings of an AAAR international specialty conference -- Particulate Matter Supersites Program and Related Studies, Atlanta, Georgia*.

McMurry, P. H. (1983). New Particle Formation In The Presence Of An Aerosol - Rates, Time Scales, And Sub-0.01 Mu-M Size Distributions. *Journal Of Colloid And Interface Science* 95, 72-80.

McMurry, P. H., and Friedlander, S. K. (1979). New Particle Formation In The Presence Of An Aerosol. *Atmospheric Environment* 13, 1635-1651.

McMurry, P. H., Wang, X., Park, K., and Ehara, K. (2002). The relationship between mass and mobility for atmospheric particles-A new technique for measuring particle density. *Aerosol Science and Technology* 36, 227-238.

McMurry, P. H., and Woo, K. S. (2002). Size distributions of 3-100-nm urban Atlanta aerosols: Measurement and observations. *Journal Of Aerosol Medicine-Deposition Clearance And Effects In The Lung* 15, 169-178.

Mebust, M. R., Eder, B. K., Binkowski, F. S., and Roselle, S. J. (2003). Models-3 community multiscale air quality (CMAQ) model aerosol component - 2. Model evaluation. *Journal Of Geophysical Research-Atmospheres* 108.

Meyer, M. B., Patashnick, H., Ambs, J. L., and Rupprecht, E. (2000). Development of a sample equilibration system for the TEOM continuous PM monitor. *Journal Of The Air & Waste Management Association* 50, 1345-1349.

Park, K., Kittelson, D. B., and McMurry, P. H. (2004a). Structural properties of diesel exhaust particles measured by transmission electron microscopy (TEM): Relationships to particle mass and mobility. *Aerosol Science And Technology* 38, 881-889.

Park, K., Kittelson, D. B., Zachariah, M. R., and McMurry, P. H. (2004b). Measurement of inherent material density of nanoparticle agglomerates. *Journal Of Nanoparticle Research* 6, 267-272.

Park, S.-K., Cobb, C. E., Wade, K., Mulholland, J., Hu, Y., and Russell, A. (2005a). Uncertainty in Air Qualiy Model Evaluation for Particulate Matter due to Spatial Variation in Pollutant Concentrations. *Atmospheric Environment, submitted for publication*.

Park, S.-K., Ke, L., B., Y., G., R. A., and Zheng, M. (2005b). Source apportionment of PM2.5 using a three-dimensional air quality model and a receptor model. *Proceedings of an AAAR international specialty conference -- Particulate Matter Supersites Program and Related Studies, Atlanta, Georgia*.

Park, S.-K., Marmur, A., Ke, L., Yan, B., Zheng, M., and Russell, A. G. (2005c). Comparison between Chamical Mass Balance Receptor and CMAQ MODEL PM2.5 Source Apportionment. *Submitted to Environmental Science and Technology*.

Park, S.-K., and Russell, A. G. (2003). Sensitivity of PM 2.5 to emissions in the Southeast. *Proceedings of the Models-3 User's workshop, Research Triangle Park, North Carolina.*

Patashnick, H., and Rupprecht, E. G. (1991). Continuous Pm-10 Measurements Using The Tapered Element Oscillating Microbalance. *Journal Of The Air & Waste Management Association* 41, 1079-1083.

Penttinen, P., Timonen, K. L., Tittanen, P., Mirme, A., Ruuskanen, J., and Pekkanen, J. (2001). Number concentration and size of particles in urban air: Effects on Spirometric lung function in adult asthmatic subjects. *Environmental Health Perspectives* 109, 319-323.

Pope, C. A., Thun, M. J., Namboodiri, M. M., Dockery, D. W., Evans, J. S., Speizer, F. E., and Heath, C. W. (1995). Particulate Air-Pollution As A Predictor Of Mortality In A Prospective-Study Of Us Adults. *American Journal Of Respiratory And Critical Care Medicine* 151, 669-674.

Pope, C. A., Verrier, R. L., Lovett, E. G., Larson, A. C., Raizenne, M. E., Kanner, R. E., Schwartz, J., Villegas, M., Gold, D. R., and Dockery, D. W. (1999). Heart rate variability associated with particulate air pollution. *American Heart Journal* 138, 890-899.

PSU/NCAR (2003). PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3. *Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research.*

Redner, A. R., and Walker, H. F. (1984). Mixture densities, maximum likelihood and the EM algorithm. *SIAM review* 26, 195-239.

Saldiva, P. H. N., Pope, C. A., Schwartz, J., Dockery, D. W., Lichtenfels, A. J., Salge, J. M., Barone, I., and Bohm, G. M. (1995). Air-Pollution And Mortality In Elderly People - A Time-Series Study In Sao-Paulo, Brazil. *Archives Of Environmental Health* 50, 159-163.

Schwab, J. J., Hogrefe, O., Demerjian, K. L., and Ambs, J. L. (2004). Laboratory characterizaton of modified tapered element oscillating microbalance samplers. *Journal Of The Air & Waste Management Association* 54, 1254-1263.

Tobias, H. J., Beving, D. E., Ziemann, P. J., Sakurai, H., Zuk, M., McMurry, P. H., Zarling, D., Waytulonis, R., and Kittelson, D. B. (2001). Chemical analysis of diesel engine nanoparticles using a nano-DMA/thermal desorption particle beam mass spectrometer. *Environmental Science & Technology* 35, 2233-2243.

Tonnesen, G. S. (2003). Regional haze modeling: Recent modeling results for VISTAS and WRAP. *Models-3 User's workshop, Research Triangle Park, North Carolina*.

Tsyro, S. (2002). First estimates of the effects of aerosol dynamics in the calculation of PM10 and PM2.5. *EMEP/MSC-W*, www.emep.int/reports/mscw_note_4_2002.pdf.

Turpin, B. J., and Lim, H. J. (2001). Species contributions to PM2.5 mass concentrations: Revisiting common assumptions for estimating organic mass. *Aerosol Science And Technology* 35, 602-610.

UCAR (2003). TDL U.S. and Canada Surface Hourly Observations. *http://dss.ucar.edu/datasets/ds472.0/*.

US-EPA (2004). SMOKE User's manual. *http://cf.unc.edu/cep/empd/products/smoke*.

Van Loy, M., Bahadori, T., and Wyzga, R. (2000). The Aerosol Research and Inhalation Epidemiology Study (ARIES): PM2.5 Mass and Aerosol Component Concentrations and Sampler Intercomparisons. *Journal of Air and Waster Management Association* 50, 1446-1458.

Vedal, S. (1997). Ambient particles and health: Lines that divide. *Journal Of The Air & Waste Management Association* 47, 551-581.

Venkataraman, C., Lyons, J. M., and Friedlander, S. K. (1994). Sire Distributions Of Polycyclic Aromatic-Hydrocarbons And Elemental Carbon.1. Sampling, Measurement Methods, And Source Characterization. *Environmental Science & Technology* 28, 555-562.

Whitby, K. T. (1978a). The physical characteristics of sulfur aerosl. *Atmospheric Environment* 12, 135-159.

Whitby, K. T. (1978b). The physical characteristics of sulfur aerosol. *Atmospheric Environment* 12, 135-159.

Wilemski, G. (1984). Composition Of The Critical Nucleus In Multicomponent Vapor Nucleation. *Journal Of Chemical Physics* 80, 1370-1372.

Woo, K. S., Chen, D. R., Pui, D. Y. H., and McMurry, P. H. (2001a). Measurement of Atlanta aerosol size distributions: Observations of ultrafine particle events. *Aerosol Science And Technology* 34, 75-87.

Woo, K. S., Chen, D. R., Pui, D. Y. H., and Wilson, W. E. (2001b). Use of continuous measurements of integral aerosol parameters to estimate particle surface area. *Aerosol Science And Technology* 34, 57-65.

Zhang, Q., Canagaratna, M. R., Jayne, J. T., Worsnop, D. R., and Jimenez, J. L. (2005). Time- and size-resolved chemical composition of submicron particles in Pittsburgh: Implications for aerosol sources and processes. *Journal Of Geophysical Research-Atmospheres* 110.

Zhang, Y., Pun, B., Vijayaraghavan, K., Wu, S. Y., Seigneur, C., Pandis, S. N., Jacobson, M. Z., Nenes, A., and Seinfeld, J. H. (2004). Development and application of the model of aerosol dynamics, reaction, ionization, and dissolution (MADRID). *Journal Of Geophysical Research-Atmospheres* 109.

CHAPTER 5

COMPARISON BETWEEN CHEMICAL MASS BALANCE RECEPTOR AND CMAQ $PM_{2.5}$ SOURCE APPORTIONMENT

(S.-K. Park, A. Marmur, L. Ke, B. Yan, M. Zheng, and A.G. Russell.

Environmental Science and Technology, in peer review)

Abstract

Source apportionment of particulate matter has been commonly performed using receptor models, but studies suggest that the assumptions in receptor models limit the accuracy of results. An alternative approach is the use of three dimensional source-oriented air quality models. Here, a comparison is done between the $PM_{2.5}$ apportioned from a receptor model with those from a source-based air quality model, and an analysis for the strengths and limitations of each model is performed. Chosen models are the Chemical Mass Balance (CMB) receptor model using organic tracers as molecular markers and the source-based Community Multiscale Air Quality (CMAQ) model. Source apportionment is conducted at sites in the southeastern United States for July 2001 and January 2002. $PM_{2.5}$ source apportionment results had moderate discrepancies, which originate from different spatial scales, fundamental limitations, and uncertainties of the two models. Results from CMB fluctuated temporally more than real variation due to measurement errors, whereas those from CMAQ could not capture daily variation well. In addition, results from CMB are mass contributions for the monitoring location, whereas those from CMAQ represent the average mass contributions of the model grid. It is difficult to assess which approach is "better". Indeed, both models have strengths and

limitations, and each model's strengths can be utilized to help overcome the other model's limitations.

5.1 Introduction

In 1997, the National Ambient Air Quality Standard (NAAQS) for $PM_{2.5}$ (particulate matter with aerodynamic diameter less than 2.5 micrometers) was promulgated in a response to scientific studies linking elevated fine particle concentrations with adverse health effects (Dockery and Pope, 1994; Metzger et al., 2004; Peel et al., 2002). Effective control of $PM_{2.5}$ levels requires identifying and quantifying contributions by major sources. Historically, source apportionment of $PM_{2.5}$ largely has been performed via several receptor-modeling techniques. In particular, one of the most widely used receptor modeling techniques is the chemical mass balance (CMB) approach (Core et al., 1982; Friedlander, 1973; Watson et al., 2002b). In CMB, ambient chemical concentrations are expressed as the sum of products of source profiles and source contributions. This linear system of equations is solved for source contributions by weighted least square fitting.

The CMB model is based, in part on the following six assumptions (US-EPA, 1990): 1. Compositions of source emissions are constant over the period of ambient and source sampling. 2. Chemical species do not react with each other (i.e., they add linearly). 3. All sources with a potential for contributing to the receptor have been identified and have had their emissions characterized. 4. The number of sources or source categories is less than or equal to the number of species. 5. Source profiles are linearly independent of each other. 6. Measurement uncertainties are random, uncorrelated, and normally distributed. Those six assumptions are never totally satisfied in actual practice (Watson et

al., 2002b), and deviations from assumptions increase the uncertainty of the source contribution estimates.

The first assumption is violated, to an unknown extent, as various source emission profiles change, for example, biomass burning (e.g., wildfires vs. burning of wood construction waste), or vehicle emissions during rush hour versus open freeways. One way to deal with the first assumption would be to develop source profiles that allow for chemical and physical transformations of source emissions prior to sample collection. In some studies, source profiles are theoretically or empirically "aged" prior to use (Chow et al., 1988; Gordon et al., 1984; Watson and Chow, 2001; Watson et al., 2002a; Watson et al., 2002b). However, these models typically contain simplified chemical mechanisms, relative transformation and/or deposition rates, mixing volumes and transport times, thus, add more assumptions.

The second assumption is violated when constituents of the $PM_{2.5}$, in particular, organic compounds or gas-phase precursors of the $PM_{2.5}$, react with each other (Limbeck et al., 2003; Paulsen et al., 2005). Thus, the second assumption limits CMB to apportion only primary species (only about half of the $PM_{2.5}$ mass) that do not chemically change appreciably during transport. The third assumption is a major cause of calculated differences in source contributions between source and receptor models (Core et al., 1982; Ryan et al., 1998; Watson and Chow, 2000; Watson et al., 2000), and the presence of unknown sources leads to substantial errors (Christensen, 2004). The fourth assumption is typically not limiting, because usually, there are many more species in the observations than source profiles (Bhave et al., 2001). The fifth assumption is violated to some degree, because many source profiles are often similar with each other (e.g.,

gasoline and diesel vehicles; biomass burning, meat cooking and vegetative detritus). Using similar profiles accounting for minor differences in source compositions leads to collinearity problems. Methods to overcome the co-linearity of source profiles include source apportionment using both gas and particle phases in source profiles (Lin et al., 1993; Marmur et al., 2005b; McKee et al., 1990; Wadden et al., 1991) or using organic compounds as tracers (Schauer et al., 1996; Zheng et al., 2002; Zheng et al., 2005). However, these approaches do not completely remove the co-linearity problem in CMB, as will be explained later. The sixth assumption is invalid, because any measurement involves not only random error but also systematic error, defined as the difference between the true value of the quantity of interest and the value to which the mean of the measurements converge as more measurements are taken (Morgan and Henrion, 2003). A further problem is that errors in source profiles and measured concentrations cause uncertainty in predicted mass contributions (Watson et al., 2002b). In addition to the above assumptions, it is inappropriate to use CMB when estimating the impact of emission reductions on species that react in a significantly non-linear fashion. Also, CMB does not link the impact of emissions to either temporal or spatial locations of the sources of $PM_{2.5}$. Nevertheless, CMB has a great advantage that results provide temporal and spatial variations at the receptor because they are based on measured concentrations.

A bottom-up approach to source apportionment is the use of emission-based, (or source) three dimensional air quality models (AQMs). Such models solve the species conservation equation:

$$\frac{\partial c_i}{\partial t} + \nabla \cdot (U\ c_i) = \nabla \rho\ D_i \nabla \left(\frac{c_i}{\rho} \right) + R_i(c_1, c_2, ..., c_n, T, t) + S_i(x,t) \qquad (1)$$

i=1, 2, 3, … , n

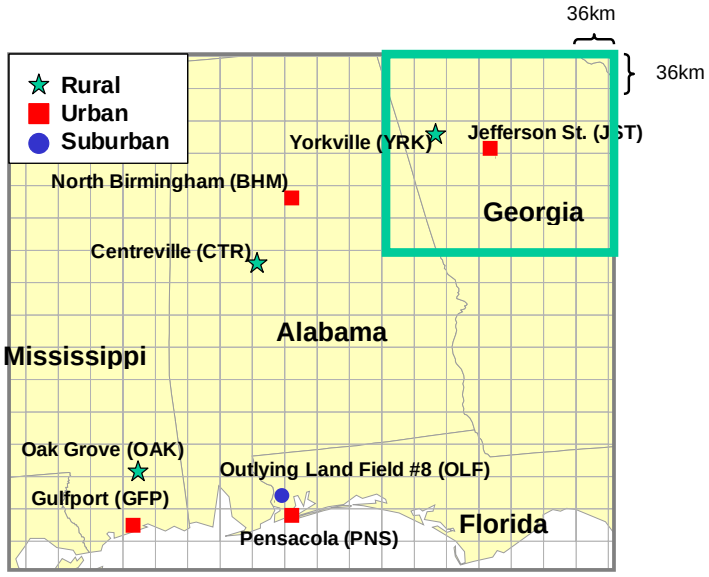
where, c_i is the concentration of species *i*, *U* is wind velocity vector, D_i is molecular diffusivity of species *i*, R_i is rate of concentration change of species *i* by chemical reaction, S_i *(x,t)* is source/sink of *i* at location *x* and time *t*, ρ is air density, and n is the number of predicted species. The conservation equation describes the formation, transport, and fate of air pollutants, including components for processing emissions, meteorology, topography, air quality observations, and chemistry (Russell and Dennis, 2000). Use of air quality models in source apportionment has limitations, principally that results are dependent on the accuracy of the model process and input data. However, such models can overcome the six assumptions of CMB. Three-dimensional models, including California Institute of Technology (CIT) photochemical model (Held et al., 2004; Kleeman and Cass, 2001; Mysliwiec and Kleeman, 2002; Ying et al., 2004), the Urban and Regional Multiscale (URM) airshed model (Boylan et al., 2002, 2005), and the Community Multiscale Air Quality (CMAQ) models (Marmur et al., 2005a; Napelenok et al., 2005; Park et al., 2005c), have been used for $PM_{2.5}$ source apportionment. Here, consistency and accuracy of source apportionment methods are compared by analyzing the discrepancies between CMB receptor and CMAQ models. Recently, source apportionment of $PM_{2.5}$ from the externally mixed source-oriented model was compared with that from the chemical mass balance (CMB) model (Held et al., 2005) in the San Joaquin Valley and in the South Coast Air Basin each for three days. Here, the

comparison was done for two complete months (July 2001 and January 2002) over eight stations in the Southeastern United States, and the reasons of discrepancy are analyzed.

5.2 Methods

Source apportionment of $PM_{2.5}$ is performed using two different approaches: Chemical Mass Balance (CMB) and Community Multiscale Air Quality (CMAQ) modeling. CMB was applied to observations at the eight Southeastern Aerosol Research and Characterization (SEARCH) stations (Figure 5-1a), and CMAQ was run over the United States (Figure 5-1b) for July 2001 and for January 2002 corresponding to the coordinated intensive monitoring periods by the US-EPA Eastern Supersite Program (ESP 01/02). SEARCH stations are located in North Birmingham, in Alabama (BHM), Centreville, Alabama (CTR), Yorkville, Georgia (YRK), Jefferson St – Atlanta, Georgia (JST), Gulfport, Mississippi (GFP), Oak Grove, Mississippi (OAK), Outlying Landing Field #8, Florida (OLF), and Pensacola, Florida (PNS) (Figure 5-1a). Daily average mass contributions of $PM_{2.5}$ were compared between the two approaches for the JST station, and monthly average results were compared at all stations.

(a)



(b)

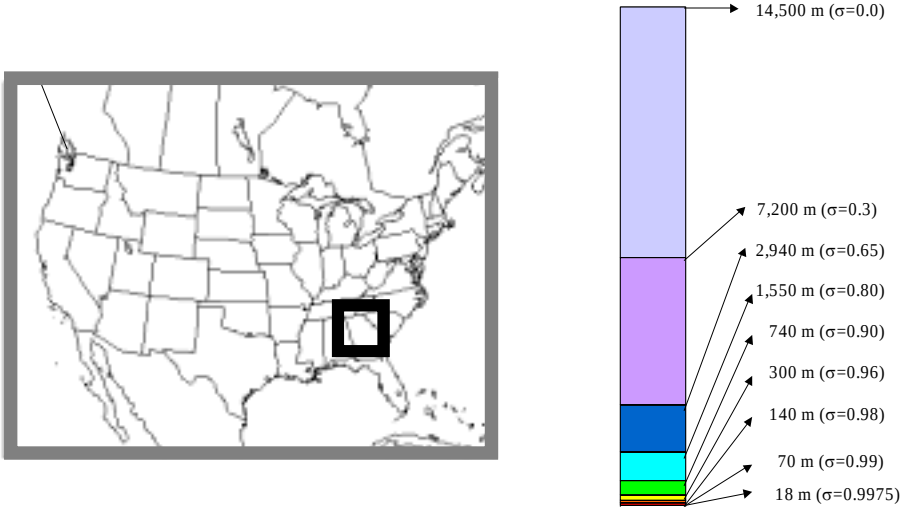


Figure 5-1. (a) SEARCH $PM_{2.5}$ monitors (http://www.atmospheric-research.com) over plotted by CMAQ 36km grid. A rectangle containing YRK and JST is CMAQ 12km domain. (b) CMAQ domain. Rectangles around the United States and over the Atlanta area are 36km and 12km grid domain, respectively. The number of vertical layer is nine with top pressure of 100hPa.

5.2.1 Receptor-based $PM_{2.5}$ apportionment using a Chemical Mass Balance (CMB) model

Receptor-based source apportionment of $PM_{2.5}$ is performed using CMB with organic tracers as Molecular Markers (CMB-MM) (Ke et al., 2005; Schauer et al., 1996; Zheng et al., 2005). Mass contributions are calculated for seven sources: gasoline exhaust, diesel exhaust, road dust, wood/biomass burning, meat cooking, natural gas, and power plant emissions. Historically, source profiles for the CMB application are composed of elemental species that include sulfate, nitrate, ammonium, elemental carbon, organic carbon, and trace metals. CMB has a long record of use, but the co-linearity of profiles relying solely on inorganic species proved problematic in application to the southeastern United States. Thus, CMB-MM, which relies more on speciated organic compounds was developed (Schauer et al., 1996; Zheng et al., 2002; Zheng et al., 2005). Source profiles are expressed as normalized values to organic carbon. Hence, CMB-MM apportions mass contributions to organic carbon, then the mass contributions to $PM_{2.5}$ is calculated by dividing by the organic carbon to $PM_{2.5}$ ratio of each source.

CMB-MM requires additional effort to analyze organic compounds present in $PM_{2.5}$, but the advantage of this method over the CMB is that the relative distribution of specific organic compounds in source emissions can provide extra means to fingerprint sources that cannot be uniquely identified by the elemental composition alone (Rogge et al., 1994; Rogge et al., 1993a, 1993b, 1993c, 1993d, 1997a, 1997b, 1998; Rogge et al., 1991; Schauer et al., 1999a, 1999b, 2001, 2002a, 2002b). For example, cholesterol is found in the meat cooking source profile. Thus, much of the co-linearity problem of source profiles can be alleviated.

5.2.2 Source-based $PM_{2.5}$ Apportionment using the Community Multiscale Air Quality (CMAQ) model

The three-dimensional air quality modeling system used for the source apportionment is EPA Models-3, which includes the Sparse Matrix Operator Kernel Emissions (SMOKE) v1.5 for emission processing (US-EPA, 2004e), the NCAR's 5th generation Mesoscale Model (MM5) v3.5.3 for preparing meteorological fields (PSU/NCAR, 2003), and the Community Multi-scale Air Quality (CMAQ) model v4.3 for air quality modeling (Byun and Ching, 1999). CMAQ was applied for July 2001 and January 2002 over the continental United States and parts of Mexico and Canada with a 36 km grid and over the Atlanta area with a 12 km grid (Figure 5-1b). The projection used is the unified Regional Planning Organization (RPO) national grid, which is Lambert conformal conic projection with a central meridian of 97 W, a center of latitude of 40 N, and standard parallels of 33 N and 45 N. More information of the air quality modeling system and the model evaluation results are available in elsewhere (Park et al., 2005b).

Source apportionment using CMAQ can be done by direct sensitivity methods such as DDM-3D (Dunker, 1981; Yang et al., 1997) or by Brute Force (BF) (i.e., applying the model once with, then without the target source) (Park et al., 2005c), the latter of which was applied in this study. The target emission source is removed based on the source category code (SCC) in the emission inventory. Emission sources apportioned were the same as the seven emission categories chosen in CMB for comparison purposes. CMAQ can calculate mass contributions to the secondary $PM_{2.5}$ in addition to those to the

primary $PM_{2.5}$. Only mass contributions to primary $PM_{2.5}$ from CMAQ were compared with those from CMB due to the limitations of CMB.

5.3 Results

Monthly $PM_{2.5}$ compositions in the Southeastern United States were apportioned using CMB-MM and CMAQ (36 km grid) (Figure 5-2). Sulfate, nitrate, and ammonium masses are not shown to give more focus on primary $PM_{2.5}$ mass. While mass contributions calculated from CMAQ and those from CMB-MM are similar, a reasonable amount of discrepancy exists. Total primary $PM_{2.5}$ masses are different between the two methods as $PM_{2.5}$ mass from CMB-MM is the measured concentration and that from CMAQ is simulated value. Relative $PM_{2.5}$ mass and contributions from each source using CMAQ do not differ significantly between July 2001 and January 2002, whereas, those from CMB-MM vary markedly from July 2001 to January 2002. Reasons of this difference will be discussed later. Daily $PM_{2.5}$ masses at the JST station apportioned using CMB-MM and CMAQ (12 km) were compared as well (Figure 5-3). CMAQ mass contributions were high on some days compared with observations, and low on other days. This daily variation would not be captured if the analysis was done only for a few days, so analysis for long periods are necessary to understand the temporal trends and the robustness of the comparison over periods comparable to air quality management decision making.

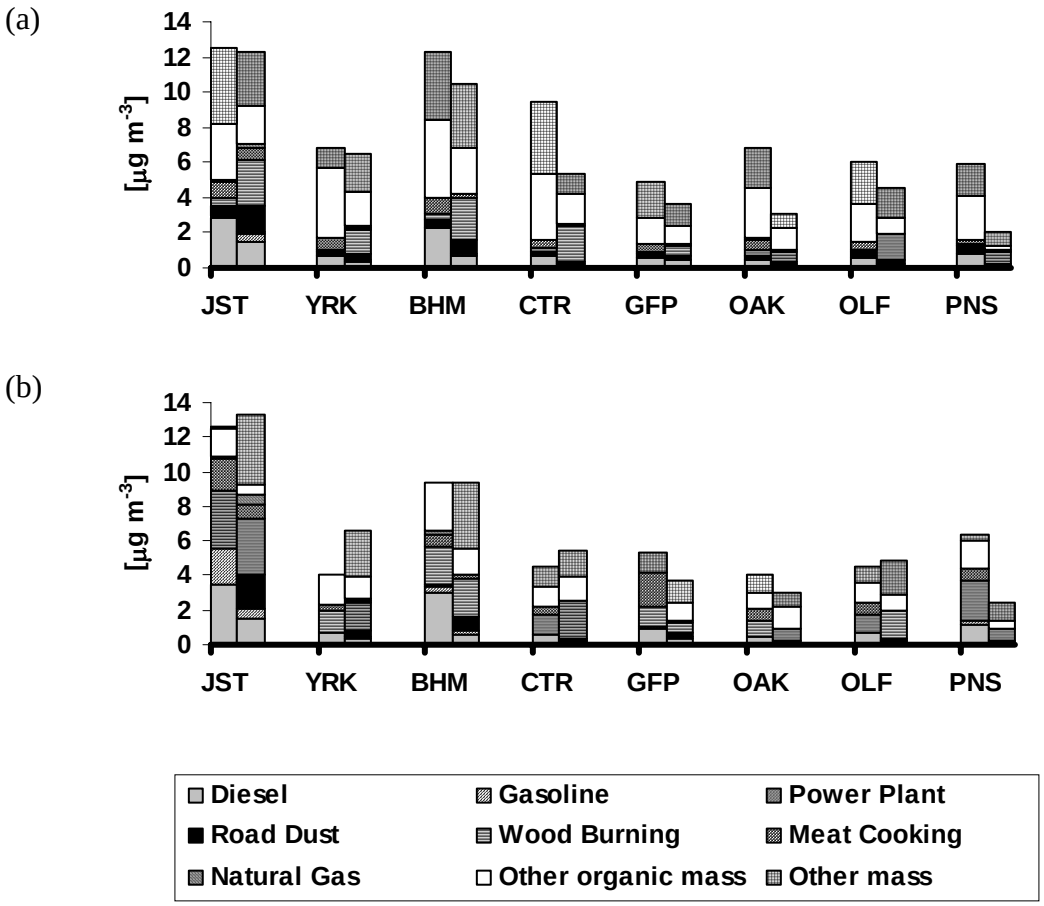


Figure 5-2. Monthly average mass contributions to $PM_{2.5}$ in SEARCH stations (sulfate, nitrate, and ammonium were excluded). CMB-MM, CMAQ (36km) [left to right]. (a) July 2001. (b) January 2002.

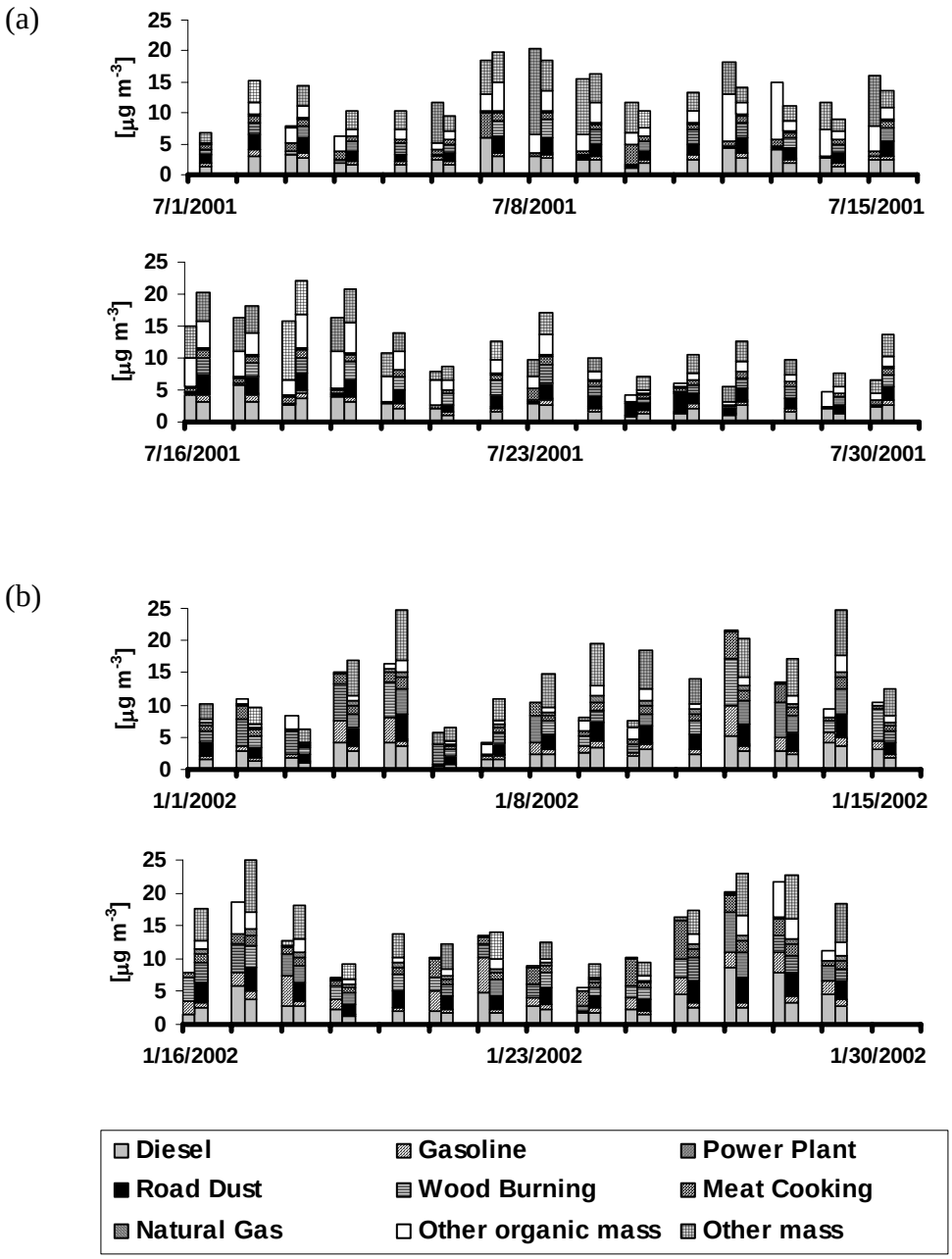


Figure 5-3. Daily average mass contributions to $PM_{2.5}$ in JST (sulfate, nitrate, and ammonium were excluded). CMB-MM, CMAQ (12km) [left to right]. Results of CMB-MM are not available on July 1, 2, 5, 22, 24, and 28, and on January 1, 11, and 20. (a) July 2001. (b) January 2002.

Monthly average mass contributions of $PM_{2.5}$ from CMAQ (36 km) and that from CMB-MM at eight SEARCH stations were positively correlated, and the correlation coefficient for monthly contribution was slightly higher than that for daily mass contributions (Figure 5-4a and 5-4b). The low correlation coefficient for daily average mass contribution is partly due to zero values of mass contributions from CMB-MM (Figure 5-4b), which was not seen in the monthly average mass contributions. The correlation for spatially and temporally averaged mass contributions was quite good (0.74), with a slope near 1, apparently hiding the real differences between CMB-MM and CMAQ results (Figure 5-4c). Therefore, the source of disagreement of the two models cannot be fully analyzed using averaged mass contributions (in space or in time), or results from a few sites or from a few days, and the analysis for multiple stations for long periods is necessary.

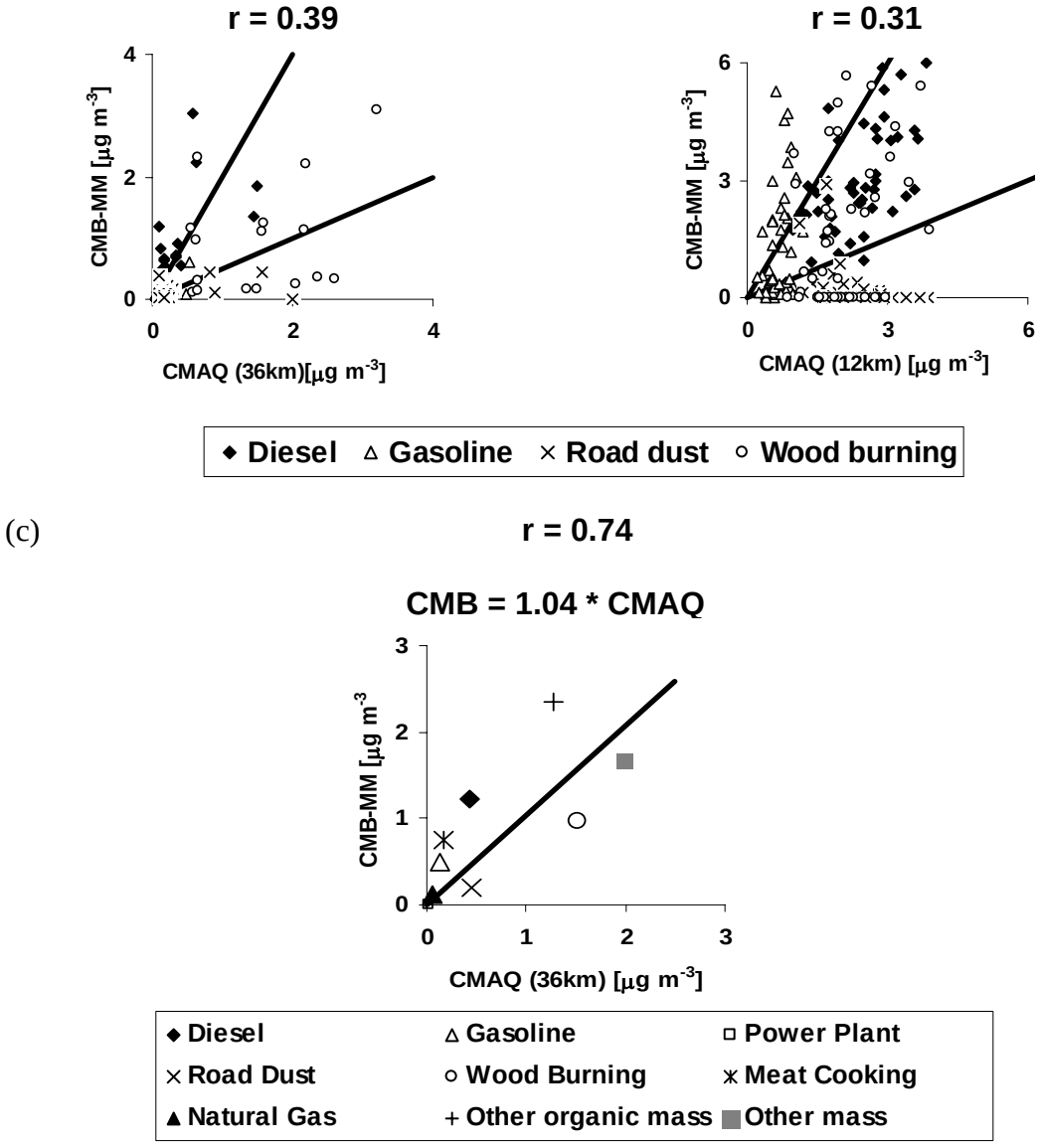


Figure 5-4. Scatter plot of mass contributions to $PM_{2.5}$ in July 2001 and January 2002. (a) Monthly average contributions in SEARCH stations. (b) Daily average contributions in JST. (c) Mass contributions averaged over the eight SEARCH stations for July 2001 and January 2002.

5.4 Discussion

Figures 5-2, 5-3, and 5-4 suggested that the separate source apportionment techniques give comparable results on average, but differences increase when considering daily source apportionments. Sources of disagreement include organic carbon to $PM_{2.5}$ ratio, different spatial and temporal variations of the two models, and uncertainties in the application of each model. To analyze the reasons of discrepancy is a fundamental to improving the accuracy of source apportionment of $PM_{2.5}$.

5.4.1 Organic Carbon to $PM_{2.5}$ ratio

The source profile used in CMB-MM is expressed as normalized values to organic carbon (OC). CMB-MM apportions mass contributions to OC, and the contributions to $PM_{2.5}$ are calculated by dividing the contributions to OC by the OC to $PM_{2.5}$ ratio (Figures 5-2, 5-5, and 5-6). CMAQ also uses a speciation profile, which includes the OC to $PM_{2.5}$ ratio, to differentiate $PM_{2.5}$ emissions to sub categories. Currently, the OC to $PM_{2.5}$ ratios used in CMB-MM and CMAQ are different (Figure 5-6). Therefore, although the mass contributions to OC from CMB-MM and CMAQ are similar, the different OC to $PM_{2.5}$ ratio can cause the mass contribution to $PM_{2.5}$ of CMB-MM to be significantly different from that of CMAQ, and vice-versa. Studies showed that the ratio varies in a wide range depending on the location or experimental situation of the emission analysis (Chow et al., 2004; Cooper, 1981; Hildemann et al., 1991; Javitz et al., 1988; Schauer et al., 1999a, 1999b, 2001, 2002a, 2002b; Watson et al., 2001). Currently, the same OC to $PM_{2.5}$ ratio for each source category was applied to all the monitors in the Southeastern United States in the source apportionment with CMB-MM.

Applying different OC to $PM_{2.5}$ ratios to different regions should be considered to improve the accuracy of CMB source apportionment.

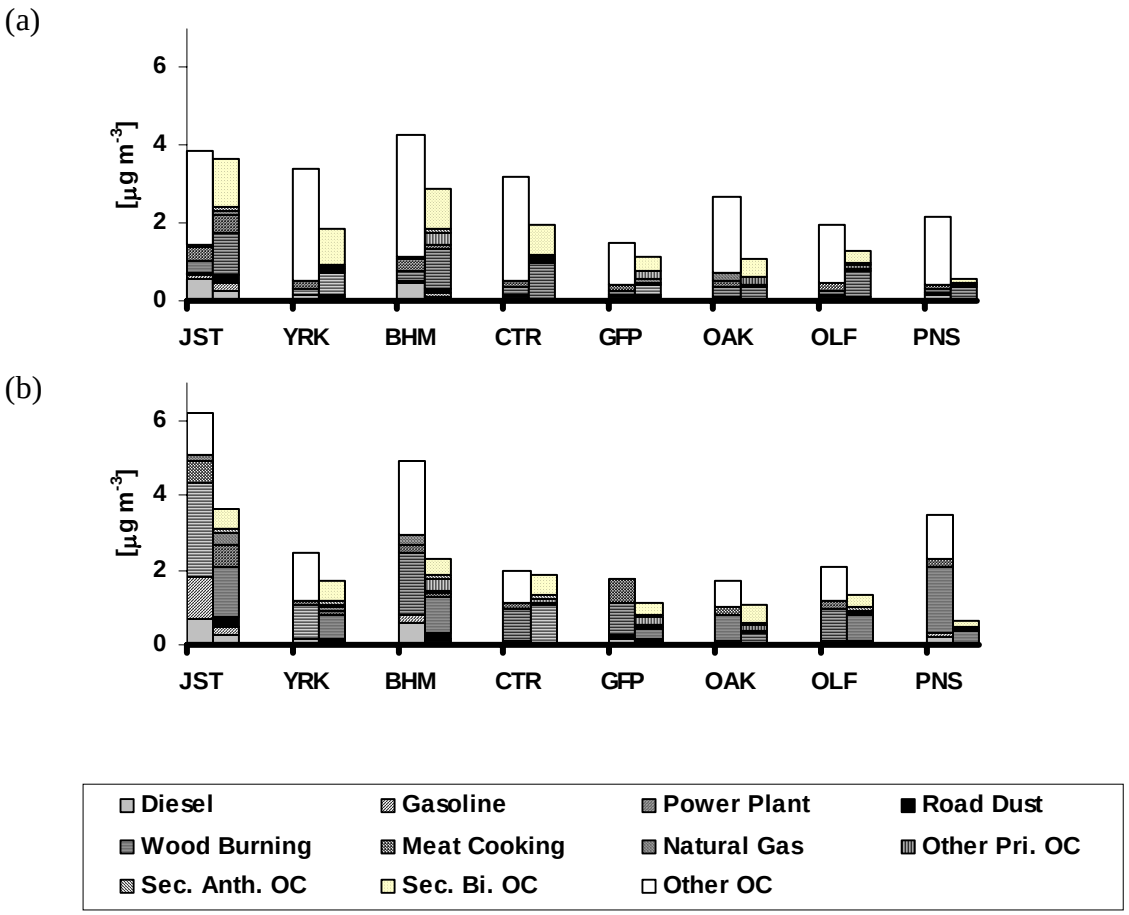


Figure 5-5. Monthly average mass contributions to organic carbon in SEARCH stations. CMB-MM, CMAQ (36km) [left to right]. (a) July 2001. (b) January 2002.

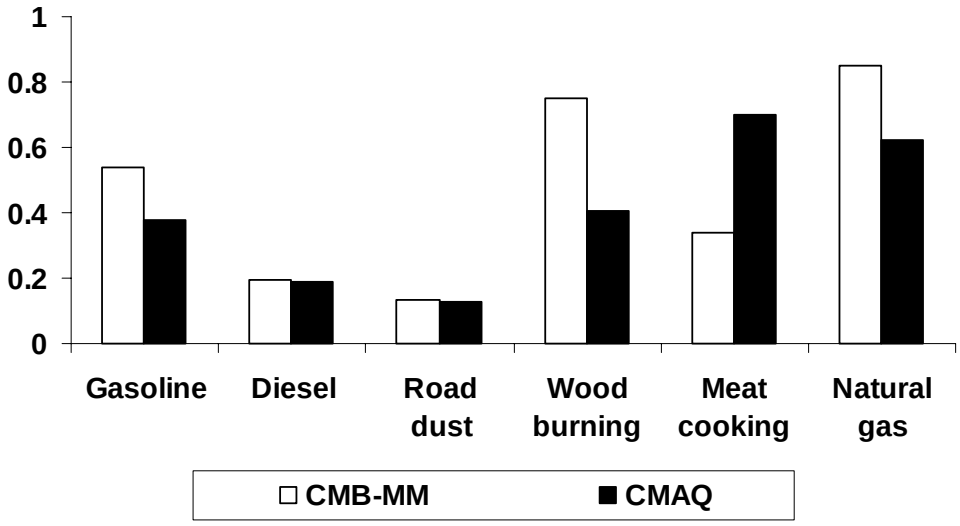


Figure 5-6. Organic carbon to $PM_{2.5}$ ratio.

5.4.2 Spatial Scale

Mass contributions calculated from CMAQ and CMB have different spatial scales. CMB calculates mass contributions of $PM_{2.5}$ at the monitoring location, whereas CMAQ simulates average concentrations of the grid, hence source contributions at the grid scale level. This different spatial scale can be a major source of the discrepancy between the two models in places where spatial gradients are relatively large. For example, primary $PM_{2.5}$ concentrations estimated from CMAQ were markedly lower than those from CMB in PNS (Figure 5-2). The PNS station is located near the Florida coast and the corresponding grid of CMAQ to PNS is occupied by more than 90% ocean (Figure 5-7). Thus, the volume-averaged concentrations are much lower than the point concentrations in PNS. Due to the large percentage of the ocean in the PNS grid, the

agreement of primary $PM_{2.5}$ mass between CMAQ for the grid containing the OLF station and CMB in the PNS station is markedly better than that between CMAQ in PNS and CMB in PNS (Figure 5-7). An additional issue is that even a minor source very near a monitor may be responsible for a large impact at that receptor, but have a small impact over a typical grid.

(a) (b)

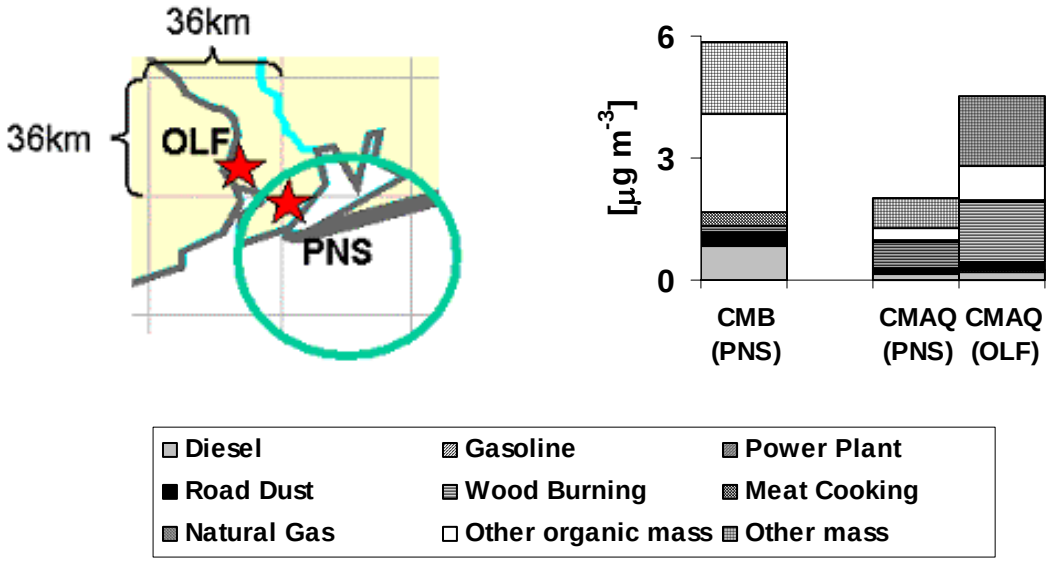


Figure 5-7. (a) Geographic location of the PNS and OLF sites over plotted by CMAQ 36km grid. (b) Monthly average mass contributions to $PM_{2.5}$ in July 2001 (re-plotted using Figure 2a).

The different spatial scales of the two models have an important implication for use of the results. CMB is done based on the measurement, so results are specific to the monitoring location. However, mass contributions calculated from CMB may not be

representative to the area if a local (spatial) concentration gradient exists. The site representativeness problem is important to epidemiological studies, because source impacts determined at the monitoring station are used to analyze the health effect of the pollutants over the area in which the monitor is located (Wade et al., 2004). On the other hand, CMAQ simulates average concentrations of the grid, so results are less subject to local effects or overestimation. However, source-based models are sensitive to errors in emissions and meteorological fields.

To build multiple monitors in the area of interests may diminish the non-representativeness (in space) of CMB results although this approach would encounter additional costs. To decrease the grid size of the air quality model can enhance the spatial resolution of CMAQ results. The mass contributions calculated from CMAQ using different sizes of the grid are compared in JST and YRK (Figure 5-8). The JST station is located in the urban area, and the YRK station is placed in the rural area. The correlation of mass contributions from CMAQ using different size of grid is lower in JST than in YRK because emissions are more localized in the urban area. In addition, mass contributions from the 36km grid are usually higher than those from 12km in YRK, suggesting that emission strength at the YRK station is weaker than that in the surrounding area. The Mean Fractional Error (MFE), Mean Fractional Bias (MFB), and correlation coefficient between mass contributions from CMB-MM and those from the two different grid sizes of CMAQ are compared (Table 5-1). The negative MFB indicates that CMAQ had lower mass contributions compared to CMB. Errors between CMB-MM and CMAQ (12km) are not always smaller than those between CMB-MM and CMAQ (36km). This result suggests that there are also significant sources other than the different

spatial scale that caused the results from CMB-MM and CMAQ to be different. This will be dealt with in detail.

Note that the advantage of the finer grid size only can be fully obtained when the resolution of emissions is also enhanced. Currently, input emissions to CMAQ were prepared for county level data that has a spatial scale on the order of 10km. Greater spatial detail is obtained by mapping emissions using surrogates, a process that cannot fully capture the spatio-temporal distributions. Thus, even though the size of the grid in the air quality model chosen is finer than 10km, the spatial resolution of the sources may not lead to significant model improvement.

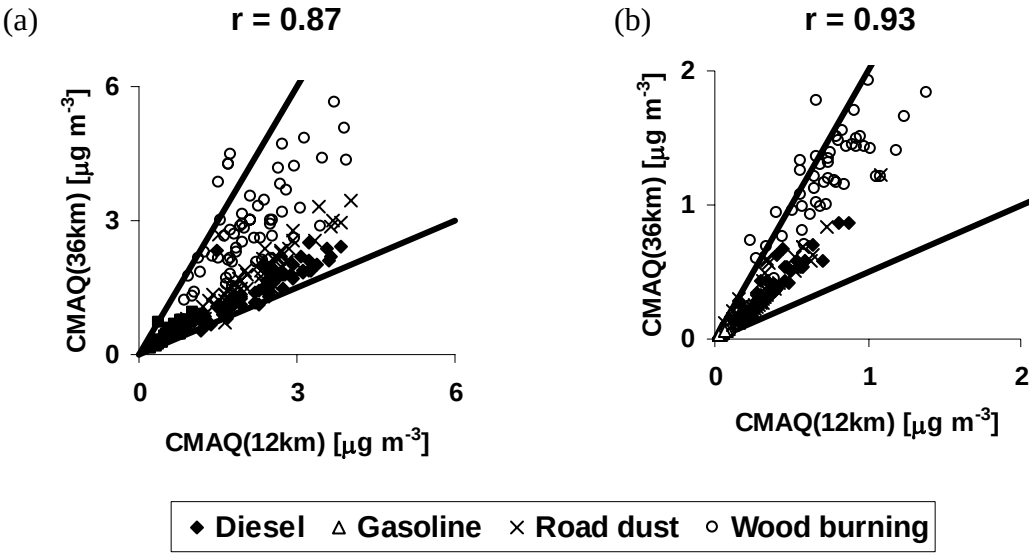


Figure 5-8. Scatter plot of daily average mass contributions to $PM_{2.5}$ from CMAQ in July 2001 and January 2002. (a) JST (urban area). (b) YRK (rural area).

Table 5-1. Mean Fractional Error (MFE), Mean Fractional Bias (MFB), and correlation (r) between daily average mass contributions in JST from CMB (MM), CMAQ (12km), and CMAQ (36km). The positive MFB indicates mass contributions from CMAQ are higher than those from CMB.

			CMB (MM) vs. CMAQ (12km)	CMB (MM) vs. CMAQ (36km)
MFE [%]	Jul. 2001	Diesel	35.8	56.9
		Gasoline	122.1	109.1
		Road dust	156.0	153.2
		Wood burning	185.5	189.6
	Jan. 2002	Diesel	41.9	70.0
		Gasoline	86.4	103.0
		Road dust	199.6	199.5
		Wood burning	50.3	55.9
MFB [%]	Jul. 2001	Diesel	-12.2	-51.6
		Gasoline	122.1	108.8
		Road dust	147.0	138.2
		Wood burning	185.5	189.6
	Jan. 2002	Diesel	-27.0	-67.0
		Gasoline	-79.2	-99.6
		Road dust	199.6	199.5
		Wood burning	-17.6	9.2
r (correl.)	Jul. 2001	Diesel	0.57	0.50
		Gasoline	0.17	0.10
		Road dust	-0.30	-0.35
		Wood burning	-0.30	-0.16
	Jan. 2002	Diesel	0.55	0.62
		Gasoline	0.35	0.30
		Road dust	-0.20	-0.18
		Wood burning	0.47	0.29

$$MFE = \frac{1}{N}\sum_{i=1}^{N}\frac{|C_1 - C_2|}{\left(\frac{C_1 + C_2}{2}\right)} \qquad MFB = \frac{1}{N}\sum_{i=1}^{N}\frac{(C_1 - C_2)}{\left(\frac{C_1 + C_2}{2}\right)}$$

Note, C_1 and C_2 are pollutant concentrations compared. N equals the number of C_1 and C_2 pairs drawn from all valid monitoring station data for the comparison time period of interest

5.4.3. Temporal Variation

Analysis showed that daily source apportionment results from CMAQ did not agree as well with those from CMB-MM (Figures 5-2 and 5-3) as did the monthly averages. Note that monthly average value has better agreement between CMB-MM and CMAQ partly because monthly mass contributions are more spatially averaged than daily mass contributions in terms of source impact. Temporal averaging also leads, effectively, to spatial averaging of source impact as winds change directions. CMB uses daily measurements, so results will capture the day-to-day temporal variation. However, the variation can fluctuate more than the true degree due to measurement errors and colinearity of source profiles in the CMB calculations. Mass contributions of road dust in January 2002 from CMB-MM are found for only one day, January 7, 2002 (Figure 5-3) because of very low levels of Al and Si in the January 2002 samples.

On the contrary, CMAQ simulates pollutant concentrations based on an annual emission inventory that is then disaggregated to account for monthly, daily, and hourly variations. Continuous emissions monitors (CEMs) on major point sources add temporal accuracy, but are limited to only the largest sources of SOx and NOx. Modeled temperature somewhat corrects mobile and biogenic emissions. It is suspected that such emissions underestimate the temporal variations of source activities, because emissions are prepared by applying typical statistical trends to the annual emission inventory. The typical trends do not include irregular events (e.g., forest fires or “bumper-to-bumper” traffic), and the trends are often smoother than the real variations because they are average values over the e.g., United States) (Figure 5-9). Thus, seasonal variations of mass contributions in addition to daily variations using CMAQ were significantly smaller

than those from CMB-MM (Figures 5-2 and 5-3), and temporal variation of CMAQ is mainly driven by the variation of meteorology due to relatively constant emissions.

(a)

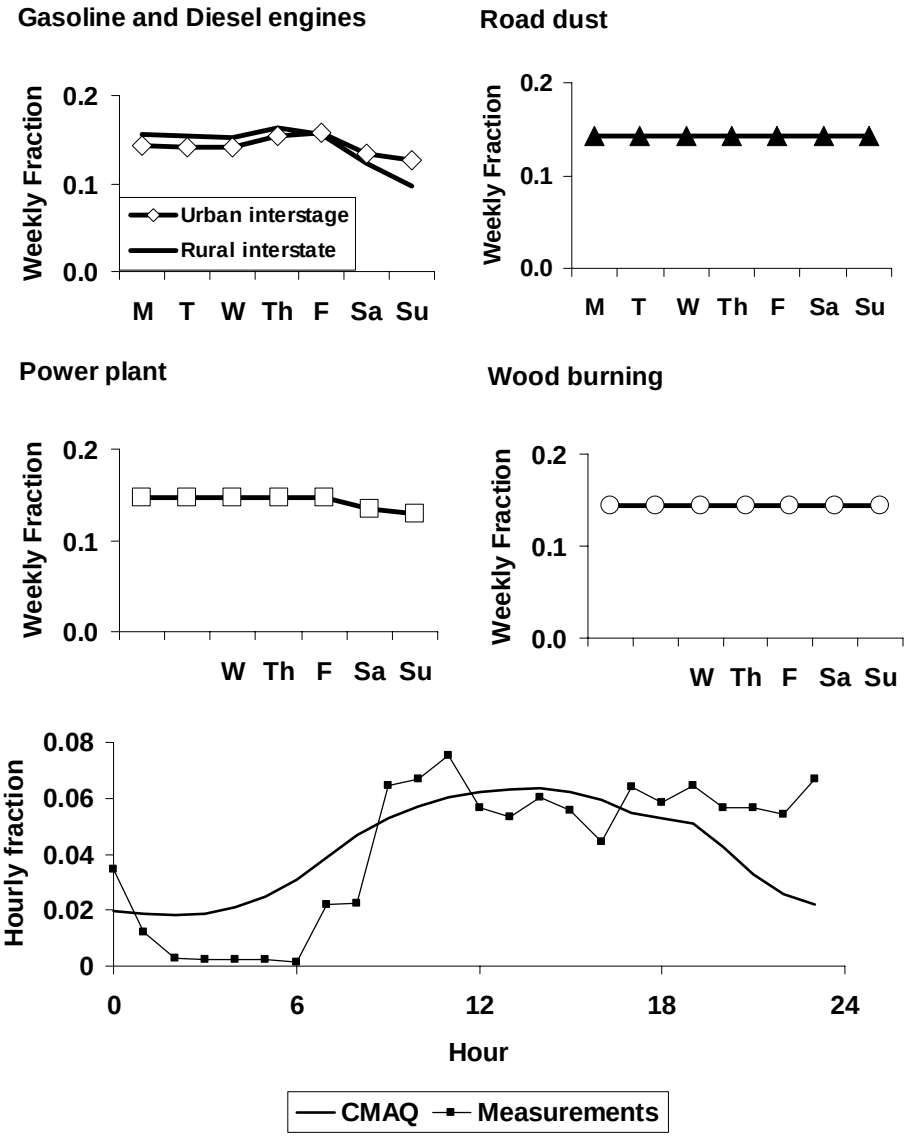


Figure 5-9. (a) Daily temporal profile of emissions in CMAQ. (b) Hourly temporal profile of commercial aircraft emissions at the Hartsfield-Jackson international airport in Atlanta, Georgia (Unal et al., 2004.).

The influence of the meteorological fields on the CMAQ results is clear from the analysis of wind speed and pollutant concentrations. All the primary pollutant concentrations from CMAQ are negatively correlated with wind speed (Figure 5-10), indicating the effect of increased dilution. The negative correlation was higher in winter than in summer because the wind speed is higher in wintertime. However, road dusts in July and in January using CMB-MM are positively correlated with wind speed (Figure 5-10). The positive correlation between road dust mass and wind speed is reasonable because more particles would be suspended and be transported more efficiently in the air when the wind is strong.

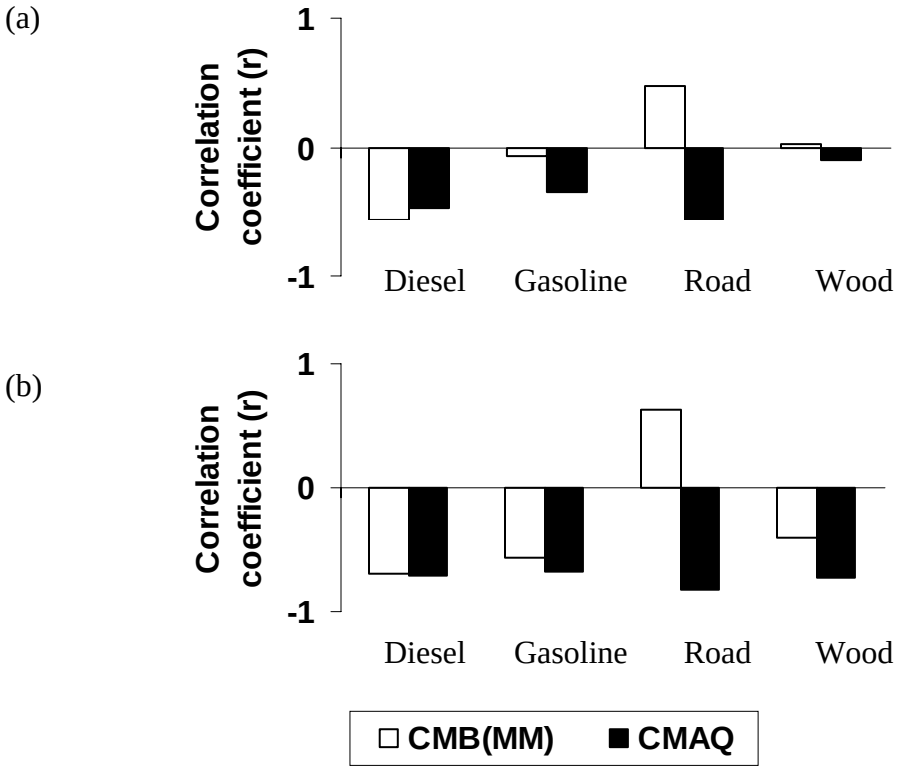


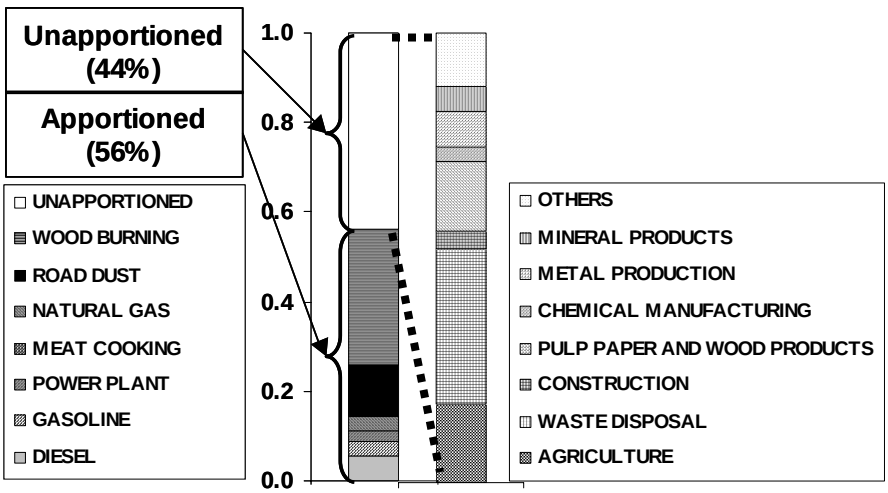
Figure 5-10. Correlation coefficient (r) between daily contributed mass to $PM_{2.5}$ and wind speed in JST. (a) July 2001. (b) January 2002.

5.4.4. Uncertainty in Each Model

Another reason for disagreement between the two models includes uncertainty of the two models. Important sources of uncertainty in CMB results include source profiles. Studies show that mass contributions estimated from CMB are significantly different depending on source profiles chosen (Yan et al., 2004). Another source of uncertainty is that CMB apportions primary mass, which is only a fraction of total $PM_{2.5}$ mass (Figures 5-2 and 5-3). Unknown sources are important not only because they occupy a large parts of $PM_{2.5}$ mass, but also because unknown sources can affect estimating known sources in CMB (Christensen, 2004). According to 1999 National Emission Inventory (1999 NEI), major $PM_{2.5}$ sources that did not have profiles for use here include agriculture, waste disposal, and wood products (Figure 5-11).

Major sources of uncertainty of CMAQ results include the emission inventory, speciation profiles, and meteorological inputs (Placet et al., 2000b). Currently, the emission inventory is known to be one of the more uncertain inputs (Abdel-Aziz and Frey, 2004; Gilliland et al., 2003; Hogrefe et al., 2003; Mannschreck et al., 2002; Mendoza-Dominguez and Russel, 2001; Placet et al., 2000b; Taghavi et al., 2005; Vautard et al., 2003). Two recent 2001 emission inventories over the southeastern United States are significantly different (Figure 5-11). Depending on sources, up to 300% of difference was observed. One method to improve the accuracy of emissions would be to calculate scaling factors of emissions via inverse modeling by incorporating measured concentrations. Indeed, both models have strengths and limitations, and each model's strength can be utilized to overcome the other model's limitations.

(a)



(b)

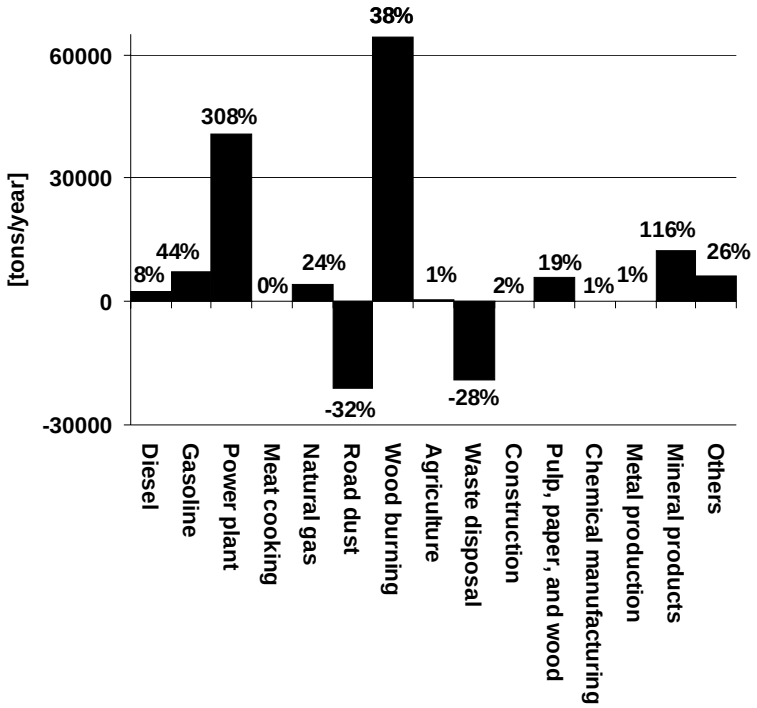


Figure 5-11. (a) $PM_{2.5}$ emissions in the southeastern US (AL, GA, FL, MS) in 2001 based on 1999 National Emission Inventory (NEI 99) (5.7E+5 tons $year^{-1}$). (b) Difference of 2001 $PM_{2.5}$ emissions in the southeastern US between emissions based on NEI 99 and those based on EPA Platform 2001 (EPA 2001 – NEI 99). The normalized difference is expressed as a percentage.

Acknowledgement

This research was supported by the U.S. Environmental Protection Agency under Agreements RD82897602, RD83107601, and RD83096001. We thank John Jansen in the Southern Company, Eric Edgerton in the Atmospheric Research & Analysis, Inc, and researchers involved in ASACA, EPA-STN, IMPROVE, and SEARCH projects for their assistance and hard work, developing the data used. We also appreciate comments from Sergey Napelenok at GA Tech.

References

Abdel-Aziz, A., and Frey, H. C. (2004). Propagation of uncertainty in hourly utility NOx emissions through a photochemical grid air quality model: A case study for the Charlotte, NC, modeling domain. *Environmental Science & Technology* **38**, 2153-2160.

Bhave, P. V., Fergenson, D. P., Prather, K. A., and Cass, G. R. (2001). Source apportionment of fine particulate matter by clustering single-particle data: Tests of receptor model accuracy. *Environmental Science & Technology* **35**, 2060-2072.

Boylan, J. W., Odman, M. T., Wilkinson, J. G., Russell, A. G., Doty, K. G., Norris, W. B., and McNider, R. T. (2002). Development of a comprehensive, multiscale "one-atmosphere" modeling system: application to the Southern Appalachian Mountains. *Atmospheric Environment* **36**, 3721-3734.

Boylan, J. W., Odman, M. T., Wilkinson, J. G., Russell, A. G., Doty, K. G., Norris, W. B., and McNider, R. T. (2005). Integrated assessment modeling of atmospheric pollutants in the Southern Appalachian Mountains. Part I: Hourly and seasonal ozone. *Journal Of The Air & Waste Management Association* **55**, 1019-1030.

Byun, D. W., and Ching, J. K. (1999). Science algorithms of the EPA Models-3 Community Multiscale Air Qualtiy (CMAQ) modeling system. *Atmospheric modeling division, National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711*.

Chow, J. C., Frazier, C., A., and Watson, J. G. (1988). A survey of existing fugitive/area source characterization methods for receptor modeling. In: Hoffnagle, G. F.; Lebowitz, M. (Eds.). *Proceedings, Particulate matter/Fugitive Dusts:*

Measurements and Control in Western Arid Regions. Air Pollution Control Association, Pittsburgh, PA., 176-189.

Chow, J. C., Watson, J. G., Kuhns, H., Etyemezian, V., Lowenthal, D. H., Crow, D., Kohl, S. D., Engelbrecht, J. P., and Green, M. C. (2004). Source profiles for industrial, mobile, and area sources in the Big Bend Regional Aerosol Visibility and Observational study. *Chemosphere* **54**, 185-208.

Christensen, W. F. (2004). Chemical mass balance analysis of air quality data when unknown pollution sources are present. *Atmospheric Environment* **38**, 4305-4317.

Cooper, J. A. (1981). Determination of Source Contributions to Fine and Coarse Suspended Particulate Levels in Petersville, Alabama. *Report to Tennessee Valley Authority by NEA, Inc.*

Core, J. E., Cooper, J. A., Hanrahan, P. L., and Cox, W. M. (1982). Particulate Dispersion Model Evaluation - A New Approach Using Receptor Models. *Journal Of The Air Pollution Control Association* **32**, 1142-1147.

Dockery, D. W., and Pope, C. A. (1994). Acute Respiratory Effects of Particulate Air-Pollution. *Annual Review of Public Health* **15**, 107-132.

Dunker, A. M. (1981). Efficient Calculation Of Sensitivity Coefficients For Complex Atmospheric Models. *Atmospheric Environment* **15**, 1155-1161.

Friedlander, S. (1973). Chemical Element Balances And Identification Of Air-Pollution Sources. *Environmental Science & Technology* **7**, 235-240.

Gilliland, A. B., Dennis, R. L., Roselle, S. J., and Pierce, T. E. (2003). Seasonal NH3 emission estimates for the eastern United States based on ammonium wet concentrations and an inverse modeling method. *Journal Of Geophysical Research-Atmospheres* **108**.

Gordon, G. E., Pierson, W. R., Daisey, J. M., Lioy, P. J., Cooper, J. A., Watson, J. G., and Cass, G. R. (1984). Considerations For Design Of Source Apportionment Studies. *Atmospheric Environment* **18**, 1567-1582.

Held, T., Ying, Q., Kaduwela, A., and Kleeman, M. (2004). Modeling particulate matter in the San Joaquin Valley with a source-oriented externally mixed three-dimensional photochemical grid model. *Atmospheric Environment* **38**, 3689-3711.

Held, T., Ying, Q., Kleeman, M. J., Schauer, J. J., and Fraser, M. P. (2005). A comparison of the UCD/CIT air quality model and the CMB source-receptor model for primary airborne particulate matter. *Atmos. Environ. accepted for publication.*

Hildemann, L. M., Markowski, G. R., and Cass, G. R. (1991). Chemical-Composition Of Emissions From Urban Sources Of Fine Organic Aerosol. *Environmental Science & Technology* **25**, 744-759.

Hogrefe, C., Sistla, G., Zalewsky, E., Hao, W., and Ku, J. Y. (2003). An assessment of the emissions inventory processing systems EMS-2001 and SMOKE in grid-based air quality models. *Journal Of The Air & Waste Management Association* **53**, 1121-1129.

Javitz, H. S., Watson, J. G., and Robinson, N. F. (1988). Performance of the chemical mass balance model with simulated local-scale aerosols. *Atmospheric Environment* **22**, 2309–2322.

Ke, L., Wang, F., Edgerton, E. S., Schauer, J. J., Russell, A. G., Cass, G. R., and Zheng, M. (2005). Source Apportionment of Organic Carbon in PM2.5 from the SEARCH study. *Atmospheric Environment, Submitted for publication*.

Kleeman, M. J., and Cass, G. R. (2001). A 3D Eulerian source-oriented model for an externally mixed aerosol. *Environmental Science & Technology* **35**, 4834-4848.

Lin, J., Scheff, P. A., and Wadden, R. A. (1993). Development of a Two-Phase Receptor Model for NMHC and PM10 Air Pollution Sources in Chicago. *Proceeding of 86th annual meeting and exhibition of the Air & Waste Management Association, Denver, Colorado*.

Mannschreck, K., Klemp, D., Kley, D., Friedrich, R., Kuhlwein, J., Wickert, B., Matuska, P., Habram, M., and Slemr, F. (2002). Evaluation of an emission inventory by comparisons of modelled and measured emission ratios of individual HCs, CO and NOx. *Atmospheric Environment* **36**, S81-S94.

Marmur, A., Park, S.-K., Mulholland, J., and Russell, A. G. (2005a). PM2.5 Source Apportionment Using Receptor and Source-oriented Models: Conceptual Differences and Implications for Health Studies. *Proceedings of an AAAR international specialty conference -- Particulate Matter Supersites Program and Related Studies, Atlanta, Georgia*.

Marmur, A., Unal, A., Russell, A. G., and Mulholland, J. A. (2005b). Optimization based source apportionment of $PM_{2.5}$: Incorporating Gas-to-Particle ratios. *Environ. Sci. Technol. accepted for publication*.

McKee, G. A., Wadden, R. A., and Scheff, P. A. (1990). Development of a Two-Phase Chemical Mass Balance Receptor Model. *Proceeding of the 83rd annual meeting and exhibition of the Air & Waste Management Association, Pittsburgh, Pennsylvania*.

Mendoza-Dominguez, A., and Russel, A. G. (2001). Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. *Atmospheric Environment* **35**, 2879-2894.

Metzger, K. B., Tolbert, P. E., Klein, M., Peel, J. L., Flanders, W. D., Todd, K., Mulholland, J. A., Ryan, P. B., and Frumkin, H. (2004). Ambient air pollution and cardiovascular emergency department visits. *Epidemiology* **15**, 46-56.

Morgan, M. G., and Henrion, M. (2003). A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. *Cambridge University Press.*

Mysliwiec, M. J., and Kleeman, M. J. (2002). Source apportionment of secondary airborne particulate matter in a polluted atmospbere. *Environmental Science & Technology* **36**, 5376-5384.

Napelenok, S., Cohan, D. S., Russell, A. G., and Hu, Y. (2005). Implementation of the direct decoupled method (DDM) into the CMAQ model for aerosol species. *Proceedings of an AAAR international specialty conference -- Particulate Matter Supersites Program and Related Studies, Atlanta, Georgia*.

Park, S.-K., Cobb, C. E., Wade, K., Mulholland, J., Hu, Y., and Russell, A. G. (2005a). Uncertainty in Air Quality Model Evaluation from Spatial Variation. *Submitted to Atmos. Environ.*

Park, S.-K., Ke, L., B., Y., G., R. A., and Zheng, M. (2005b). Source apportionment of PM2.5 using a three-dimensional air quality model and a receptor model. *Proceedings of an AAAR international specialty conference -- Particulate Matter Supersites Program and Related Studies, Atlanta, Georgia*.

Peel, J., Tolbert, P., Klein, M., Metzger, K., Flanders, W. D., Todd, K., Mulholland, J., Ryan, P. B., and Frumkin, H. (2002). Ambient air pollution and respiratory emergency department visits in Atlanta, August 1998-August 2000 (ARIES/SOPHIA). *Epidemiology* **13**, S124-S124.

Placet, M., Mann, C. O., Gilbert, R. O., and Niefer, M. J. (2000). Emissions of ozone precursors from stationary sources: a critical review. *Atmos. Environ.* **34**, 2183-2204.

PSU/NCAR (2003). PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3. *Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research.*

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., and Cass, G. R. (1994). Sources Of Fine Organic Aerosol.6. Cigarette-Smoke In The Urban Atmosphere. *Environmental Science & Technology* **28**, 1375-1388.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1993a). Sources Of Fine Organic Aerosol.2. Noncatalyst And Catalyst-Equipped Automobiles And Heavy-Duty Diesel Trucks. *Environmental Science & Technology* **27**, 636-651.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1993b). Sources Of Fine Organic Aerosol.3. Road Dust, Tire Debris, And Organometallic Brake Lining Dust - Roads As Sources And Sinks. *Environmental Science & Technology* **27**, 1892-1904.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1993c). Sources Of Fine Organic Aerosol.4. Particulate Abrasion Products From Leaf Surfaces Of Urban Plants. *Environmental Science & Technology* **27**, 2700-2711.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1993d). Sources Of Fine Organic Aerosol.5. Natural-Gas Home Appliances. *Environmental Science & Technology* **27**, 2736-2744.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1997a). Sources of fine organic aerosol.7. Hot asphalt roofing tar pot fumes. *Environmental Science & Technology* **31**, 2726-2730.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1997b). Sources of fine organic aerosol.8. Boilers burning No. 2 distillate fuel oil. *Environmental Science & Technology* **31**, 2731-2737.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1998). Sources of fine organic aerosol. 9. Pine, oak and synthetic log combustion in residential fireplaces. *Environmental Science & Technology* **32**, 13-22.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simonelt, B. R. T. (1991). Sources Of Fine Organic Aerosol.1. Charbroilers And Meat Cooking Operations. *Environmental Science & Technology* **25**, 1112-1125.

Russell, A., and Dennis, R. (2000). NARSTO critical review of photochemical models and modeling. *Atmospheric Environment* **34**, 2283-2324.

Ryan, W. M., West, C. R., Holtz, D. R., Peters, T. A., Cooper, J. A., and Ono, D. M. (1998). Reconciliation of receptor and dispersion modeling impacts of PM10 in Hayden, Arizona, In: Mathai, C. V.; Stonefield, D. H. (Eds.), Transactions, PM10: Implementation of Standards. *Air Pollution Control Association, Pittsburgh, PA.*

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T. (1999a). Measurement of emissions from air pollution sources. 1. C-1 through C-29 organic compounds from meat charbroiling. *Environmental Science & Technology* **33**, 1566-1577.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T. (1999b). Measurement of emissions from air pollution sources. 2. C-1 through C-30 organic compounds from medium duty diesel trucks. *Environmental Science & Technology* **33**, 1578-1587.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T. (2001). Measurement of emissions from air pollution sources. 3. C-1-C-29 organic compounds from fireplace combustion of wood. *Environmental Science & Technology* **35**, 1716-1728.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T. (2002a). Measurement of emissions from air pollution sources. 4. C-1-C-27 organic compounds from cooking with seed oils. *Environmental Science & Technology* **36**, 567-575.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T. (2002b). Measurement of emissions from air pollution sources. 5. C-1-C-32 organic compounds from gasoline-powered motor vehicles. *Environmental Science & Technology* **36**, 1169-1180.

Schauer, J. J., Rogge, W. F., Hildemann, L. M., Mazurek, M. A., and Cass, G. R. (1996). Source apportionment of airborne particulate matter using organic compounds as tracers. *Atmospheric Environment* **30**, 3837-3855.

Taghavi, M., Cautenet, S., and Arteta, J. (2005). Impact of a highly detailed emission inventory on modeling accuracy. *Atmospheric Research* **74**, 65-88.

Unal, A., Odman, T., and Russell, T. (2004.). Assessing the Impacts of Hartsfield-Jackson Atlanta International Airport on Particulate Matter and Ozone in Atlanta Metropolitan Area. *Prepared For National Aeronautics and Space Administration. 2004.*

US-EPA (1990). CMB8 User's Mannual. *Office of Air Quality, Planning and Standards, Research Triangle Park, NC 27711*.

US-EPA (2004). SMOKE User's manual. *http://cf.unc.edu/cep/empd/products/smoke*.

Vautard, R., Martin, D., Beekmann, M., Drobinski, P., Friedrich, R., Jaubertie, A., Kley, D., Lattuati, M., Moral, P., Neininger, B., and Theloke, J. (2003). Paris emission inventory diagnostics from ESQUIF airborne measurements and a chemistry transport model. *Journal Of Geophysical Research-Atmospheres* **108**.

Wadden, R. A., Scheff, P. A., Lin, J., Lee, H., Keil, C., Graf-Teterycz, J., Keehan, K., Kenshki, D., Milz, S., Holsen, T. M., and Khalili, N. (1991). Two Phase Receptor Modeling. *Proceeding of the 84th annual meeting and exhibition of the Air & Waste Management Association, Vancouver, British Columbia*.

Wade, K., Marmur, A., Mulholland, J., Russell, A., Peel, J., and Klein, M. (2004). Analyses of air quality data for an emergency department study, 1993-2002. *Epidemiology* **15**, S61-S61.

Watson, J. G., and Chow, J. C. (2000). Reconciling urban fugitive dust emissions inventory and ambient source contribution estimates: Summary of current knowledge and needed research. Report No. 6110.4D2. *Prepared for US*

Environmental Protection Agency, Research Triangle Park, NC, by Desert Research Institute, Reno, NV.

Watson, J. G., and Chow, J. C. (2001). Source characterization of major emission sources in the Imperial and Mexicali Valleys along the US/Mexico border. *Science Of The Total Environment* **276**, 33-47.

Watson, J. G., Chow, J. C., and Houck, J. E. (2001). PM2.5 chemical source profiles for vehicle exhaust, vegetative burning, geological material, and coal burning in Northwestern Colorado during 1995. *Chemosphere* **43**, 1141-1151.

Watson, J. G., Chow, J. C., Lowenthal, D. H., Robinson, N. F., Cahill, C. F., and Blumenthal, D. L. (2002a). Simulating changes in source profiles from coal-fired power stations: Use in chemical mass balance of PM2.5 in the Mount Zirkel Wilderness. *Energy & Fuels* **16**, 311-324.

Watson, J. G., Chow, J. C., and Pace, T. G. (2000). Fugitive dust emissions, In: Davis, W. T. (Ed.). *Air Pollution Engineering Manual, second ed. John Wiley, New York*, 117-135.

Watson, J. G., Zhu, T., Chow, J. C., Engelbrecht, J., Fujita, E. M., and Wilson, W. E. (2002b). Receptor modeling application framework for particle source apportionment. *Chemosphere* **49**, 1093-1136.

Yan, B., Zheng, M., and Russell, A. (2004). Uncertainty and Sensitivity Analysis of Chemical Mass Balance Modeling Using Organic Tracers for $PM_{2.5}$ Source Apportionments. *Proceeding of the AAAR 04 annual conference, Atlanta, GA.*

Yang, Y. J., Wilkinson, J. G., and Russell, A. G. (1997). Fast, direct sensitivity analysis of multidimensional photochemical models. *Environmental Science & Technology* **31**, 2859-2868.

Ying, Q., Mysliwiec, M., and Kleeman, M. J. (2004). Source apportionment of visibility impairment using a three-dimensional source-oriented air quality model. *Environmental Science & Technology* **38**, 1089-1101.

Zheng, M., Cass, G. R., Schauer, J. J., and Edgerton, E. S. (2002). Source apportionment of PM2.5 in the southeastern United States using solvent-extractable organic compounds as tracers. *Environmental Science & Technology* **36**, 2361-2371.

Zheng, M., Ke, L., Wang, F., Cass, G. R., Schauer, J. J., Edgerton, E. S., and Russell, A. G. (2005). Source apportionment of PM2.5 and their daily variations at Jefferson Street, Atlanta GA during Summer and Winter. *Journal of Air and Waste Association. Submitted for publication*.

CHAPTER 6

REGIONAL ADJUSTMENT OF EMISSION STRENGTHS VIA FOUR-DIMENSIONAL DATA ASSIMILATION

(S.-K. Park, S. Napelenok, and A.G. Russell.

Journal of Air and Waste Management Association, in preparation)

Abstract

An iterative procedure of direct sensitivity analysis and an inverse modeling technique can improve the emission estimates. This study uses four-Dimensional Data Assimilation (FDDA) to evaluate source emission strengths over the United States. The continental US is divided into six regions, and data assimilation is performed separately for each region in July 2001 and January 2002. A separate scaling factor is calculated for weekdays and weekends to capture the day-of-week variation in the emission bias. Adjustment is conducted for ten emission sources: CO (total), NH3 (total), SO_2 (total), NO_x (area/mobile/nonroad), NO_x (point), VOC (area/mobile/nonroad/point), VOC (biogenic), POA (total), PEC (total), and PMFINE (total). Results show that base emissions for CO and SO_2 sources are relatively accurate. Base emissions for PEC source are overestimated 100 %, but those for POA source are underestimated up to 70% when compared with the adjusted emissions. Emissions for NH_3, NO_x, and PMFINE sources are relatively accurate in July 2001, but those in January 2002 are around 100% higher than the adjusted emissions. Adjusted VOC emissions in July 2001 are similar to the base-case emissions, but those in January 2002 are underestimated up to 70% when compared with the adjusted emissions. The emission adjustment improves the air quality

model performance. For example, the mean fractional bias for each pollutant decreased on average 25%, suggesting that the inverse modeling successfully decreases the systematic bias residing in the emissions. Differences in the emission scaling factors between weekdays and weekends were minor in most cases.

6.1. Introduction

Photochemical air quality models are essential tools for the investigation of air quality management. Confidence in these models is bolstered when the pollutant concentrations simulated by the models compare well to observations. Accuracy of modeling results relies on model inputs, and one of the most uncertain inputs to air quality models includes the emission inventory (Placet et al., 2000b). Four Dimensional Data Assimilation (FDDA) (Mendoza-Dominguez and Russell, 2000) was developed to refine and evaluate emission inventories. The FDDA method links formal direct sensitivity analysis of three dimensional air quality models with inverse modeling, and incorporates observational data of multiple species. FDDA can estimate the emission strength of various source categories in an inventory simultaneously. It can also evaluate the process in the air quality model. This method has been assessed by applying the perturbation to the base emission inventories to generate predictions that serve as observations (or pseudo-observations) in the assimilation process, and has successfully identified the predefined perturbation applied (Mendoza-Dominguez and Russell, 2000).

Fundamental assumptions of FDDA indicate that the major source for the discrepancies between the simulated and observed concentrations of pollutants is the emission inventory, and spatial allocation of emissions are less important to overall

performance than the total mass of emissions (Mendoza-Dominguez and Russell, 2001b). These limitations have been discussed in previous studies (Mendoza-Dominguez and Russell, 2000, 2001a, 2001b). Previously, FDDA was applied to estimate adjustments in the emissions of gas-phase primary species and O_3 precursors during August 9-10, 1992 over the Atlanta, Georgia Metropolitan area (Mendoza-Dominguez and Russell, 2001b). In that study, the California Institute of Technology (CIT) airshed model (Harley et al., 1993) and ridge regressions (Frank and Friedman, 1993) were used as the forward and the inverse models, respectively. This method was also applied to derive emission adjustments of domain- wide sources of fine organic aerosol and gas-phase species (e.g., NO_x, Volatile Organic Compounds: VOCs, CO, SO_2, and NH_3) over the eastern United States during July 9-19, 1995 and May 22-29, 1995 (Mendoza-Dominguez and Russell, 2001a). The study employed the Urban-to-Regional multiscale Model (URM) (Odman and Russell, 1991) as a forward model.

The present study assessed the feasibility of using the FDDA approach to estimate regional emission strength adjustments over the continental United States. Comparing with the previous studies that were limited to the Atlanta, Georgia, metropolitan area or the eastern United States, this study extended the spatial domain over the continental United States. Also, regionally different emission scaling factors were estimated to capture spatial variations of systematic biases in emissions. Furthermore, modeling was performed for two complete months (i.e., July 2001 and January 2002). This ensures that the estimated emission adjustment better reflects systematic biases because emission-scaling factors calculated are less dependent on the short-period abnormal meteorology or modeling system deficiencies.

6.2. Methods

The FDDA approach adjusts emissions by incorporating the surface observations into the three-dimensional air quality model coordinates. The difference between observations and simulated concentrations, along with the sensitivity of pollutant concentrations to emissions, are used to estimate how much emissions from a specific source should be altered to optimize model performance. The FDDA approach is iterative in nature as the sensitivity changes with emission strengths. Detailed descriptions of FDDA are available elsewhere (Mendoza-Dominguez and Russell, 2000, 2001a, 2001b), and summary is provided below.

Let O_{ik} be the k^{th} observation (in space and time) of the i^{th} species, and $P_{ik}^{<adjusted>}$, be the adjusted prediction at the same location and time with the measured concentrations, then the weighted error is defined as:

$$e_{ik} = w_{ik} (O_{ik} - P_{ik}^{<adjusted>}), \tag{6-1}$$

where w_{ik} is a weighting factor that accounts for measurement uncertainties, and other properties of the measurements. The value of the adjusted simulated concentration (i.e., $P_{ik}^{<adjusted>}$) is approximated by the linear combination of the product of the sensitivity coefficients and the perturbed emission strengths:

$$P_{ik}^{<adjusted>} = P_{ik} + \sum_{j=1}^{J} \delta_{ij} \; s_{ik,j} \; m_j \quad , \tag{6-2}$$

where P_{ik} is the predicted concentration at base emission, $s_{ik,j}$ is the sensitivity of the i^{th} species to emissions from the j^{th} source at the k^{th} location, δ_{ij} is a scaling factor applied to the sensitivity coefficients, J is the total number of sources involved, and m_j is the emission adjustment scaling factor of the j^{th} source from its base level. The receptor model is written using equations (6-1) and (6-2) as:

$$w_{ik}\, d_{ik} = w_{ik} \sum_{j=1}^{J} \delta_{ij}\ s_{ik,j}\ m_j + e_{ik} \quad , \tag{6-3}$$

or in matrix notation:

$$\mathbf{W_e d = W_e G m + e}, \tag{6-4}$$

where $\mathbf{W_e}$ is a ($I \cdot K$ x $I \cdot K$) diagonal matrix of observation weights (the w_{ik}'s are the elements of the diagonal), **d** ($d_{ik}=O_{ik}-P_{ik}$) and **e** (error term) are ($I \cdot K$) length vectors, **G** ($g_{ik,j} = \delta_{ij} \cdot s_{ik,j}$) is the ($I \cdot K$ x J) matrix of weighted sensitivity coefficients, and **m** is a J-length vector. I, J, and K are the total number of species, sources, and monitoring stations, respectively. Weighting factors for the sensitivity coefficients (δ_{ij}) are used to over-relax the emission adjustments of those sources that are suspect of being highly uncertain and are likely to change the most during the assimilation process. Each element of the diagonal matrix $\mathbf{W_e}$, w_{ik} is computed as:

$$w_{ik} = w_{ik}^{<n>} \cdot w_{ik}^{<c>} , \tag{6-5}$$

where $w_{ik}^{<n>}$ is defined as the inverse of the total (valid) number of observations for species i, i.e., $1/N_i$. This allows each species to have equal weight in the solution. $w_{ik}^{<c>}$ is defined as the inverse of the variance of the k^{th} observation (space and time) for species i, i.e., $1/(\sigma_{ik})^2$. The more accurate observations will have greater weight than the inaccurate observations. The number of observations and the variance for each species are discussed below. Incorporation of these weighting factors in FDDA speeds up the convergence of the solution, but it has little impact on the final results (Mendoza-Dominguez and Russell, 2000).

The unknown parameter, **m**, the amount of emissions to be adjusted is selected by minimizing the following:

$$\Gamma = \mathbf{e}^T\mathbf{W_e}\mathbf{e} \; . \qquad (6\text{-}6)$$

The resulting m after minimizing equation (6-6) is expressed as:

$$\mathbf{m} = (\mathbf{G}^T\mathbf{W_e}\mathbf{G})^{-1}\mathbf{G}^T\mathbf{W_e}\mathbf{d} \; . \qquad (6\text{-}7)$$

The above formulation implies that the weighted difference between observations and predictions completely drives the emission adjustment, **m**. However, these estimates should be in a certain condition that the new emissions should lie inside a range defined by uncertainty limits of the base emission estimates. Uncertainties in emission estimates are often represented by a log-normal distribution. Assuming that the base inventory is located at the median of the distribution, a Confidence Interval (CI) of the emissions estimates ranges from $1/\sigma_g^n$ to σ_g^n (σ_g is the geometric standard deviation), where n depends on the CI selected. Here, *n* was selected as 3.3 for CI to be 99.9% (Mendoza-Dominguez and Russell, 2000). To control the emission adjustment within prescribed bounds, a penalty function is introduced. The penalty function is a (JxJ) diagonal matrix, $\mathbf{W_m}$, which has parameters w_{jj} (positive constants) in the diagonal. Note that when this matrix is incorporated into the linear system in equation (6-6), the method is modified from being observationally driven to a mixed (observation and emission) approach.

The new objective function incorporating the penalty function is given by:

$$\Gamma = \mathbf{e}^T\mathbf{W_e}\mathbf{e} + \mathbf{m}^T\mathbf{W_m}\mathbf{m} \; . \qquad (6\text{-}8)$$

The least-square estimator derived from minimizing the equation (6-8) is as follows:

$$\mathbf{m} = (\mathbf{G}^T\mathbf{W_e}\mathbf{G} + \mathbf{W_m})^{-1}\mathbf{G}^T\mathbf{W_e}\mathbf{d} \; . \qquad (6\text{-}9)$$

When all elements of $\mathbf{W_m}$ are identical, Hoerl and Kennard's Ridge Regression (RR) estimator is obtained (i.e., $\mathbf{W_m}=\lambda\mathbf{I}$, where λ is known as the ridge parameter and **I** is the identity matrix) (Hoerl and Kennard, 1970). Equation (6-9) is known as a "biased"

estimator, but it can give more stable predictions than ordinary least squares when there are few data or when data have noise (Aldrin, 1997; Frank and Friedman, 1993; Marquardt, 1963). To terminate the iteration process the following expression is used (Hoerl and Kennard, 1970):

$$\frac{\lambda^{t} - \lambda^{t-1}}{\lambda^{t-1}} \leq 20 \left\{ \frac{Tr[(G^{T} W_{e} G)^{-1}]}{J} \right\}^{-1.3} \quad \text{(for t>1)}, \qquad \text{(6-10)}$$

where λ^t is the λ (ridge parameter) at t^{th} iteration, and Tr[**M**] denotes the trace of the square matrix **M**. The error bounds (1σ) for each scaling factor are obtained from the square root of the diagonal elements of the variance-covariance matrix (**V**) in the emission adjustment estimates.

$$\mathbf{V} = (\mathbf{G^T W_e G} + \mathbf{W_m})^{-1}(\mathbf{G^T W_e G})(\mathbf{G^T W_e G} + \mathbf{W_m}) \qquad \text{(6-11)}$$

Since iteration is conducted to obtain the ultimate estimates for vector **m**, the values of **G** and $\mathbf{W_m}$ from the last iteration are used to compute **V** (Menke, 1989). For the nonlinear problem, σ does not necessarily represent a standard deviation from a typical Gaussian distribution.

Here, FDDA was performed using the US-EPA Models-3 system and ridge regression as the forward and inverse models, respectively. Ridge regression was described above, while Models-3 is composed of the Community Multiscale Air Quality (CMAQ) model for air quality modeling (Byun and Ching, 1999), the PSU/NCAR's 5th generation Mesoscale Model v5 (MM5) for meteorological modeling (PSU/NCAR, 2003), and the Sparse Matrix Operator Kernel Emissions (SMOKE) for emission processing (US-EPA, 2004e). Air quality modeling was conducted over the continental United States for July 2001 and January 2002, which correspond to the coordinated

intensive Eastern Supersite Program (ESP 01/02) period. More information of the modeling domain, parameters used in modeling, and the model evaluation results are available in Chapter 3.

6.3. Results and discussion

The continental United States is divided into six regions based on the U.S. Census Bureau (Census, 2005): Pacific (P), Mountain (M), midWest (W), Southeast (S), Northeast (N), and Georgia (G) (Figure 6-1). Emission adjustment scaling factors are calculated separately for those six regions. In addition to the different scaling factors for each region, scaling factors were calculated for weekdays (Monday through Friday) and weekends (Saturday and Sunday) to account for the day-of-week variation.

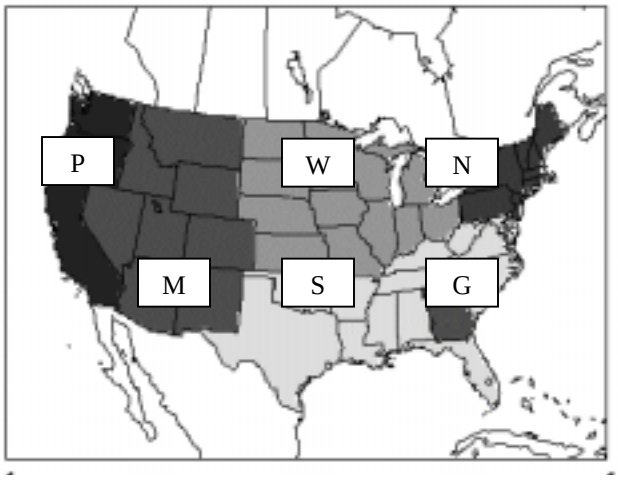


Figure 6-1. Six regions defined for the FDDA approach.

Twenty-eight species are used in the assimilation process, and the coefficient of variation for each species is illustrated in Table 6-1. The coefficient of variation multiplied by the average concentration of each species becomes the standard deviation of each species, which is used to calculate the observation weight ($w_{ik}^{<c>}$). Ten emission sources: CO (T), NH_3 (T), NO_x (AMN), NO_x (P), SO_2 (T), VOC (AMNP), VOC (B), primary organic aerosol (POA; T), primary elemental carbon (PEC; T), and crustal element (PMFINE; T), are adjusted for each region. Weighting factors for the sensitivity coefficients (δ_{ij}) are defined based on the assumption that given the actual concentration of the simulated primary species, the simulated concentration of secondary species will also be approximately correct; i.e., given the right mixture of NO_x and VOC, simulated ozone will be close to the observed value. Thus, δ_{ij} was set to 1.0 for primary species (species *i* come from source *j*). δ_{ij} was set to 0.5 for most secondary species, and δ_{ij} was set to 0.1 to give less weight because of the lack of measurements for volatile organic carbons. Some species were not assimilated because they result in severe instabilities in the adjusted estimates (Mendoza-Dominguez and Russell, 2001a). Weighting factors of the sensitivity coefficients used in this study are given in Table 6-2. The upper limits of the CI of emissions, i.e., $\sigma_g^{3.3}$, are 3.73 for VOC and POA, and 2.25 for all other sources (Mendoza-Dominguez and Russell, 2001a).

Table 6-1. Observed species assimilated and the coefficient of variation used for each species in the ridge regression module (Mendoza-Dominguez and Russell, 2001a).

Species	Coefficient of variation (%)
O_3	30
CO	50
NO_2	45
NO_x	45
NO_y	45
SO_2	40
Formaldehyde (HCHO)	50
Acetaldehyde (CCHO)	50
Acetone	50
Ethylene	50
Isoprene	50
Ethane (ALK1)	50
ALK2	50
ALK3	50
ALK4	50
ALK5	50
ARO1	50
ARO2	50
OLE1	50
OLE2	50
TNMOC	50
SO_4^{2}- $PM_{2.5}$	30
NO_3^{-} $PM_{2.5}$	50
NH_4^{+} $PM_{2.5}$	30
OC $PM_{2.5}$	40
EC $PM_{2.5}$	40
Crustal $PM_{2.5}$	60
Total $PM_{2.5}$	30

Table 6-2. Weighting factors for sensitivity coefficients (δ_{ij}) (Mendoza-Dominguez and Russell, 2001a)

Species	Sources*									
	CO (T)	NH3 (T)	NOx (AMN)	NOx (P)	SO2 (T)	VOC (AMNP)	VOC (B)	POA (T)	PEC (T)	PMFINE (T)
O_3	0.5	0.5	1	1	0.5	1	1	0.5	0	0
CO	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0	0
NO_2	0	0	0.5	0.5	0	0	0	0	0	0
NO_x	0.5	0.5	1	1	0.5	1	1	0.5	0	0
NO_y	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0	0
SO_2	0.5	0.5	0.5	0.5	1	0.5	0.5	0.5	0	0
Formaldehyde (HCHO)	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
Acetaldehyde (CCHO)	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
Acetone	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
Ethylene	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
Isoprene	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
Ethane (ALK1)	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ALK2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ALK3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ALK4	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ALK5	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ARO1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ARO2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
OLE1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
OLE2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
TNMOC	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
SO_4^{2-} $PM_{2.5}$	0	0.5	0.1	0.1	1	0	0	0.5	0	0
NO_3^- $PM_{2.5}$	0	0.5	0	0	0.5	0	0	0.5	0	0
NH_4^+ $PM_{2.5}$	0	1	0.1	0.1	0.5	0	0	0.5	0	0
OC $PM_{2.5}$	0	0.5	0.1	0.1	0.5	0	0	1	0	0
EC $PM_{2.5}$	0	0.5	0.1	0.1	0.5	0	0	0.5	1	0
Crustal $PM_{2.5}$	0	0.5	0.1	0.1	0.5	0	0	0.5	0	1
Total $PM_{2.5}$	0	0.5	0.1	0.1	0.5	0	0	0.5	0	0

* T (total): area, biogenic, mobile, nonroad and point sources
* A: area source
* B: biogenic source
* M: mobile source
* N: nonroad source
* P: point source

Emissions were adjusted separately for each region as well as for weekdays and weekends (Figures 6-2 and 6-3). The error bound (expressed as 1σ) for each scaling factor is also represented (Tables 6-3 and 6-4). Model performance with base level emissions and that with adjusted emissions for each region is illustrated in Appendix. CO emissions are relatively accurate with little difference in emission scaling factors between weekdays and weekends. Only scaling factors in the Southeast region in July 2001 and the Mountain region in January 2002 are significantly different from 1.0. These relatively small errors in the CO emissions are consistent with the previous studies (Muller and Stavrakou, 2005; Petron et al., 2004). However, a detailed comparison of the scaling factor for each region in the United States cannot be done because those studies calculated one emission-scaling factor for the North America.

NH_3 emissions are slightly overestimated in July 2001, but they are overestimated 100% in January 2002. The adjustment of NH_3 emissions improves the prediction of the nitrate pollutant concentrations, which are usually overestimated in January 2002 (Gilliland et al., 2003). Emission-scaling factors for NO_x from point sources and NO_x from area/mobile/nonroad sources had similar trends. The base NO_x emissions were overestimated around 120% for Georgia, Mountain, Northeast, and midWest regions. However, NO_x emissions were underestimated in Southeast and Pacific regions. Previous studies revealed that the NO_x in the Southeastern region was underestimated (Mendoza-Dominguez and Russell, 2001a). Also, when the global emission adjustment was conducted for the United States, the emissions decreased 14% (Muller and Stavrakou, 2005).

Emission scaling factors of PEC carbon were less than 1.0 in most cases suggesting the overestimation of the PEC emissions. The low scaling factor of PEC reflects the overestimation of the elemental carbon concentrations because the weighting factor for sensitivity coefficient (δ_{ij}) for PEC is positive only for elemental carbon species (Table 6-2). PMFINE emissions were underestimated in July 2001, but overestimated in January 2002. These results are also directly related with the performance of the crustal $PM_{2.5}$ concentrations because δ_{ij} for PMFINE is positive only for crustal $PM_{2.5}$. Studies showed that the modeling of crustal $PM_{2.5}$ has difficulty due to the uncertainty in fugitive dust emissions. Zero to 100% of the fugitive dust emissions are emitted very near ground level where it can immediately be redeposited on the ground or trapped in vegetation depending on the regions (Pace, 2003). Here, 75% of the fugitive dust emissions are assumed that they are not mixed in the modeling grid.

SO_2 emissions were relatively accurate by dint of the continuous monitoring emissions (US-EPA, 2003b). POA emissions were overestimated in most regions. This result is partly due to the underestimation of the secondary organic carbon concentrations. VOC emissions were overestimated in most cases, consistent with the previous study (Chang et al., 1996). VOC emission adjustments involved fifteen VOC species, O_3, and other species. Some VOC species concentrations were higher than the measured values, but others did not. These different model performances of multiple VOC species complicated the VOC emission adjustment.

The air quality model was run using the adjusted emissions. The performance of the model using base level emissions was compared with that using adjusted emissions. The mean fractional bias (MFB) decreased on average 25% after the adjustment (Figure

6-4), suggesting that the FDDA approach helped to minimize the systematic bias of the emissions. However, the mean fractional error (MFE) did not improve, indicating that there are other sources of error (e.g., the speciation profile of VOC), which cannot be reduced by emission adjustments alone.

The importance of the emission adjustment has been discussed in previous studies. When NO_x and VOC emissions are adjusted, the ratio of the VOC to NO_x can change from 2 (NO_x scavenges O_3) to 11 (O_3 production is favored by NO_x) (Chang et al., 1996). Thus, the emission control strategy need to be modified with the changes of the sensitivity of O_3 to NO_x emissions after the adjustment (Mendoza-Dominguez and Russell, 2001b). Also, the relative importance of the contributed mass to $PM_{2.5}$ changes as the emission adjustment is performed. For example, if the PEC emissions are adjusted negatively and POA emissions are adjusted positively, the resulting mass contributions from diesel exhaust (one of the large source of PEC emissions) would decrease, whereas those from meat cooking combustion (one of the large source of POA emissions) would increase. Therefore, accurate emissions increase the reliability of the air quality model, and thus enhance the confidence of further applications

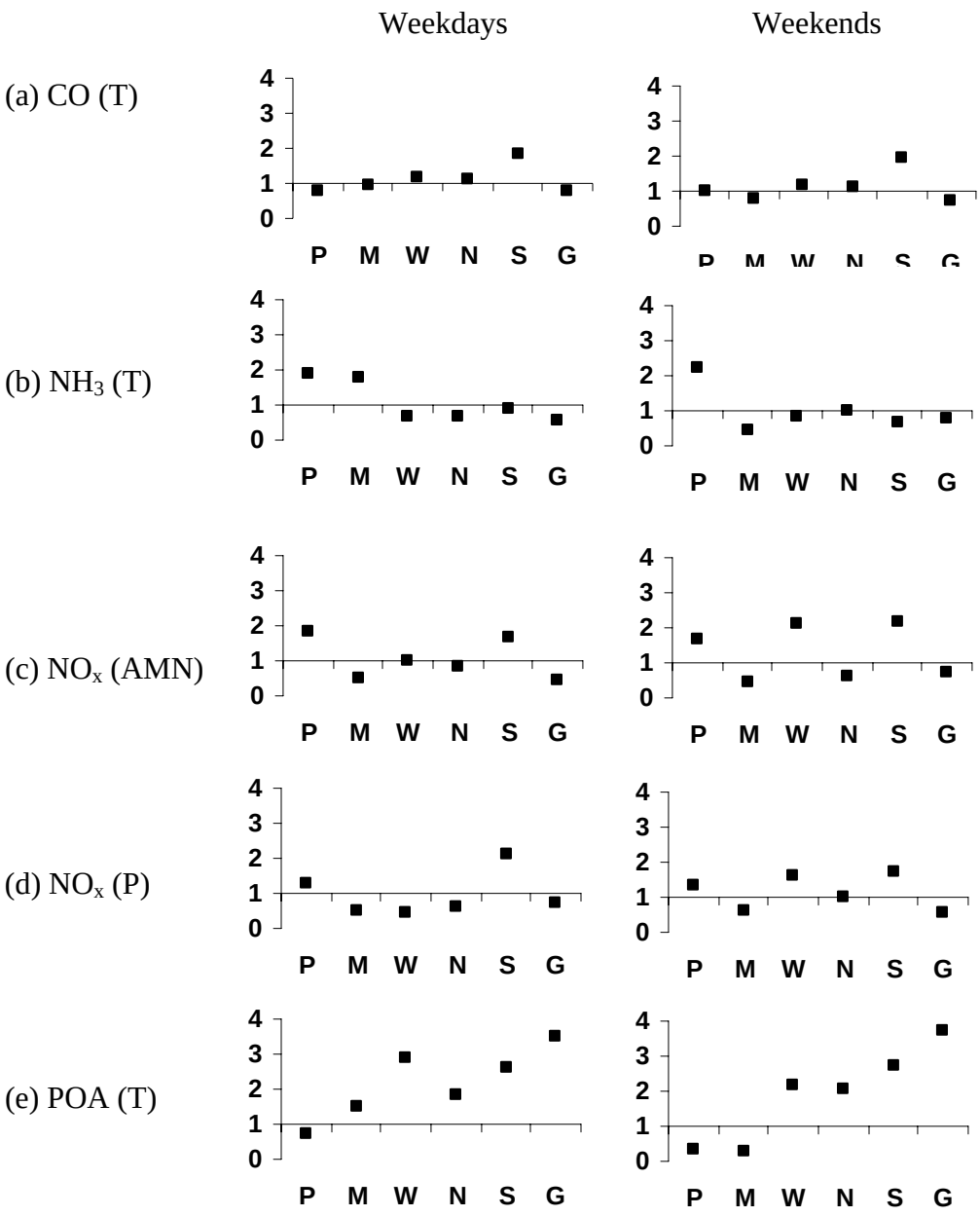


Figure 6-2. Emission adjustment scaling factors for July 2001 in the Pacific (P), Mountain (M), midWestern (W), Northeast (N), South (S), and Georgia (G). A scaling factor of one indicates that the adjusted emissions are the same as the base emissions. Two and 0.5 indicate that the adjusted emissions are twice and a half of the base emissions, respectively.

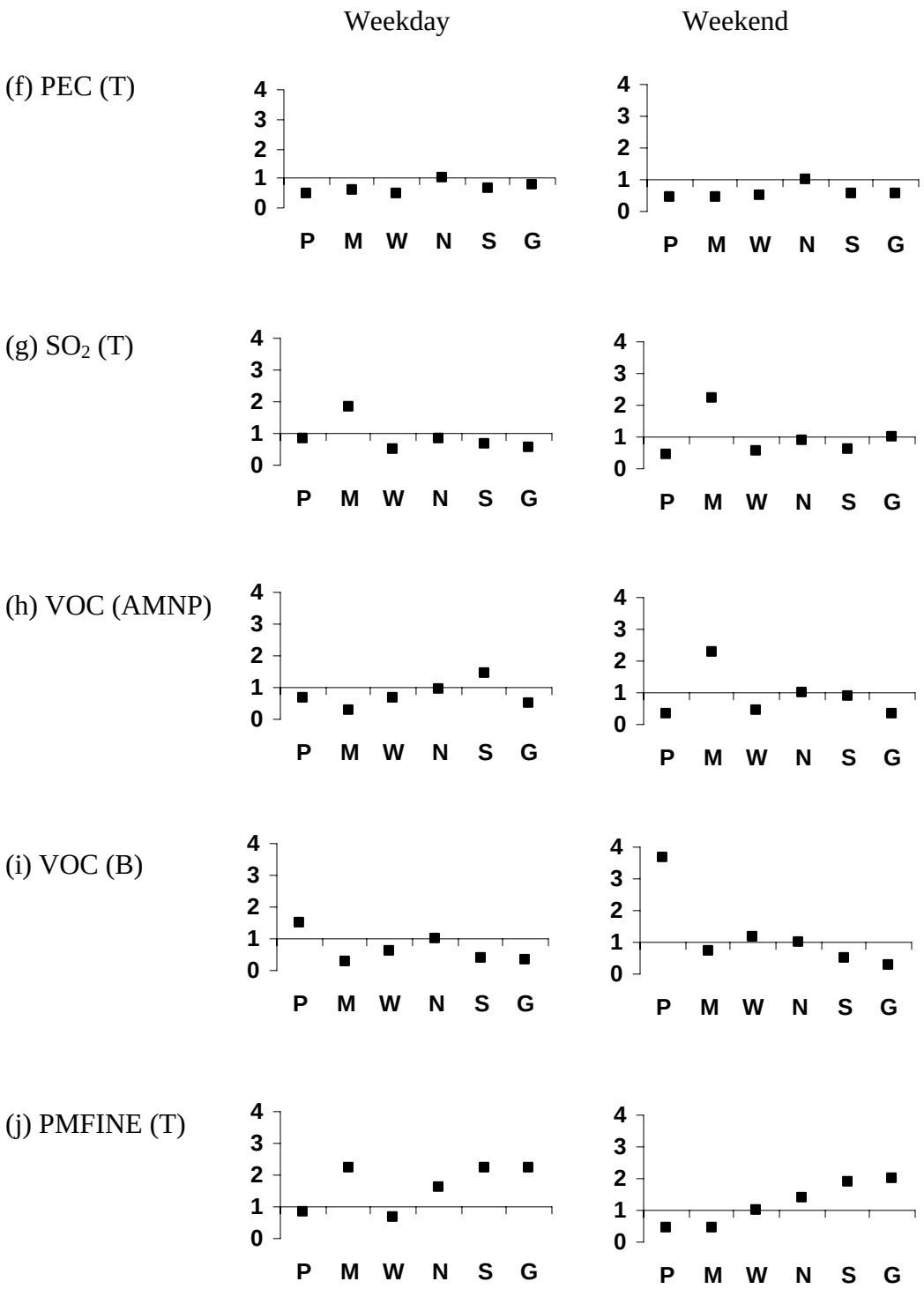


Figure 6-2. continued.

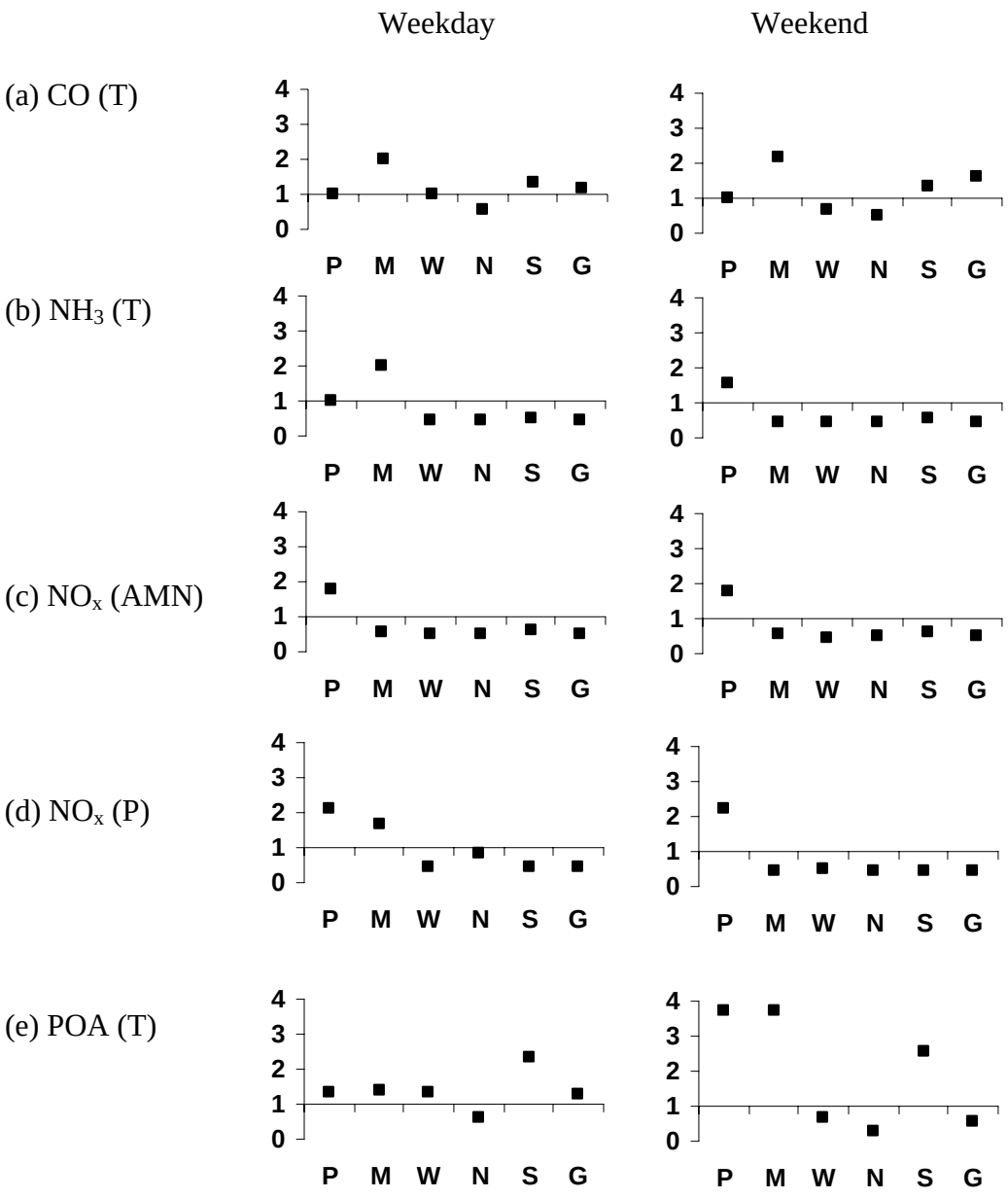


Figure 6-3. Emission adjustment scaling factor for January 2002 in the Pacific (P), Mountain (M), midWestern (W), Northeast (N), South (S), and Georgia (G). A scaling factor of one indicates that the adjusted emissions are the same as the base emissions. Two and 0.5 indicate that the adjusted emissions are twice and a half of the base emissions, respectively.

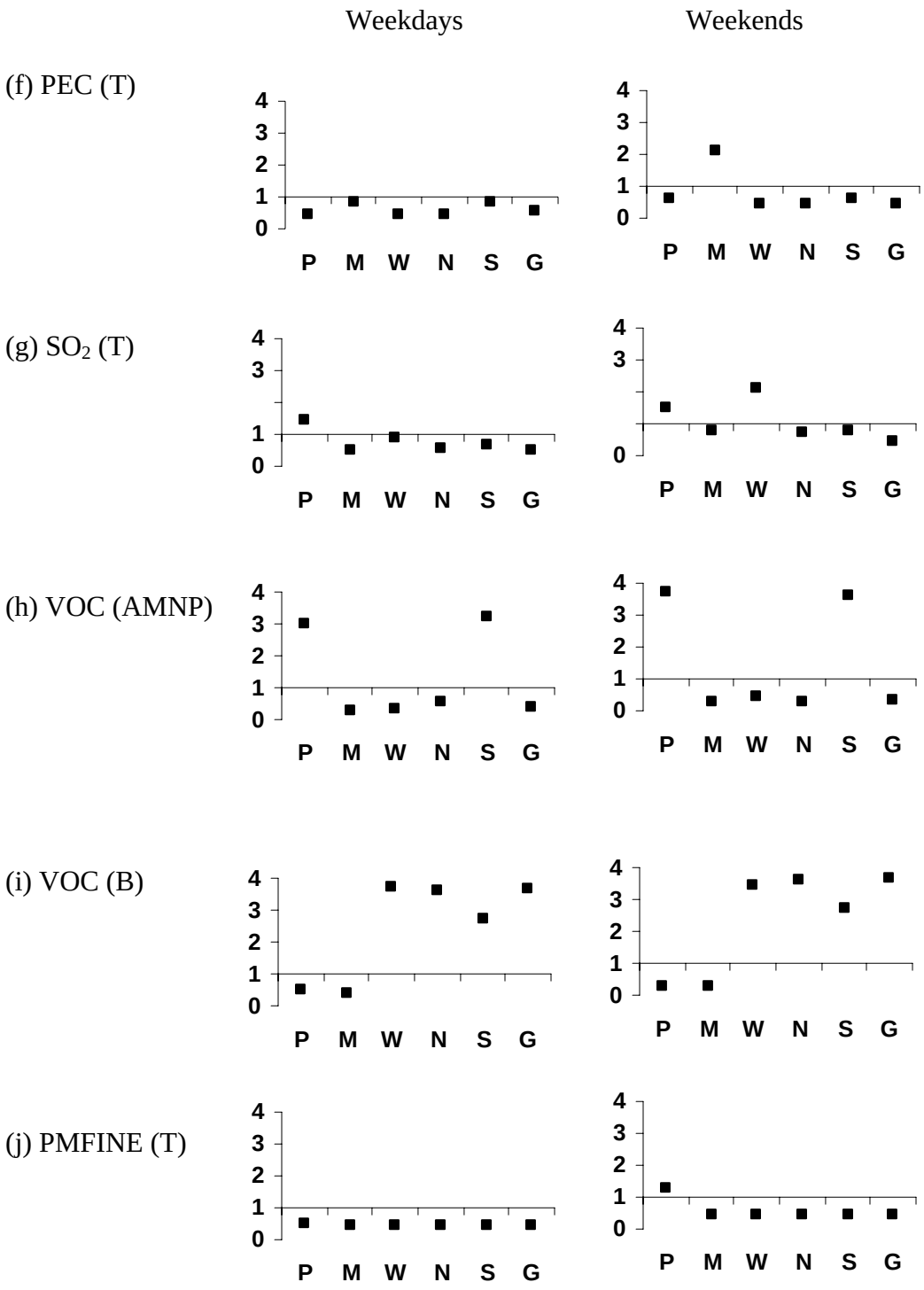


Figure 6-3. continued.

Table 6-3. Emission adjustment scaling factors and the error bound (1σ) for July 2001 in each region in (a) weekdays and (b) weekends. A scaling factor of one indicates that the adjusted emissions are the same as the base emissions. Two and 0.5 indicate that the adjusted emissions are twice and a half of the base emissions, respectively.

(a) Weekdays (Monday through Friday)

	Georgia (G)	Mountain (M)	Northeast (N)	Pacific (P)	Southeast (S)	midWest (W)
CO	0.76 ± 0.79	0.96 ± 1.14	1.11 ± 0.00	0.78 ± 1.01	1.82 ± 0.29	1.14 ± 1.09
NH3	0.58 ± 0.00	1.78 ± 0.02	0.65 ± 0.00	1.90 ± 0.00	0.87 ± 0.00	0.65 ± 0.00
NOXA	0.47 ± 0.00	0.50 ± 0.00	0.85 ± 0.00	1.86 ± 0.09	1.65 ± 0.07	0.98 ± 0.03
NOXP	0.75 ± 0.01	0.49 ± 0.07	0.60 ± 0.00	1.28 ± 0.02	2.11 ± 0.00	0.46 ± 0.00
PEC	0.78 ± 0.00	0.57 ± 0.01	1.00 ± 0.00	0.47 ± 0.00	0.65 ± 0.00	0.50 ± 0.00
PMFINE	2.24 ± 0.00	2.21 ± 0.01	1.62 ± 0.00	0.86 ± 0.01	2.23 ± 0.01	0.69 ± 0.01
POA	3.50 ± 0.00	1.51 ± 0.01	1.84 ± 0.00	0.73 ± 0.01	2.59 ± 0.00	2.90 ± 0.00
SO2	0.57 ± 0.00	1.82 ± 0.01	0.81 ± 0.00	0.85 ± 0.01	0.64 ± 0.00	0.51 ± 0.00
VOCA	0.48 ± 0.27	0.28 ± 0.01	0.97 ± 0.00	0.67 ± 0.09	1.44 ± 0.49	0.64 ± 0.08
VOCB	0.35 ± 0.32	0.28 ± 0.00	0.98 ± 0.00	1.48 ± 1.68	0.39 ± 0.57	0.63 ± 0.30

(b) Weekends (Saturday and Sunday)

	Georgia (G)	Mountain (M)	Northeast (N)	Pacific (P)	Southeast (S)	midWest (W)
CO	0.73 ± 0.76	0.77 ± 0.98	1.09 ± 0.00	0.97 ± 1.20	1.97 ± 0.19	1.14 ± 1.22
NH3	0.77 ± 0.00	0.45 ± 0.00	1.00 ± 0.00	2.20 ± 0.01	0.65 ± 0.00	0.82 ± 0.00
NOXA	0.74 ± 0.01	0.45 ± 0.00	0.61 ± 0.00	1.69 ± 0.01	2.18 ± 0.12	2.11 ± 0.00
NOXP	0.57 ± 0.01	0.59 ± 0.00	0.99 ± 0.00	1.31 ± 0.01	1.70 ± 0.01	1.60 ± 0.02
PEC	0.58 ± 0.00	0.47 ± 0.01	1.00 ± 0.00	0.45 ± 0.00	0.58 ± 0.00	0.48 ± 0.00
PMFINE	2.00 ± 0.00	0.45 ± 0.01	1.40 ± 0.00	0.45 ± 0.00	1.87 ± 0.01	1.01 ± 0.01
POA	3.72 ± 0.00	0.27 ± 0.00	2.06 ± 0.00	0.33 ± 0.02	2.70 ± 0.00	2.17 ± 0.00
SO2	0.98 ± 0.00	2.22 ± 0.01	0.91 ± 0.00	0.45 ± 0.01	0.62 ± 0.00	0.54 ± 0.00
VOCA	0.34 ± 0.03	2.28 ± 0.14	0.97 ± 0.00	0.34 ± 0.22	0.88 ± 0.47	0.42 ± 0.47
VOCB	0.31 ± 0.07	0.74 ± 0.29	0.98 ± 0.00	3.65 ± 0.00	0.51 ± 0.48	1.15 ± 0.24

Table 6-4. The same as Table 6-3 for January 2002

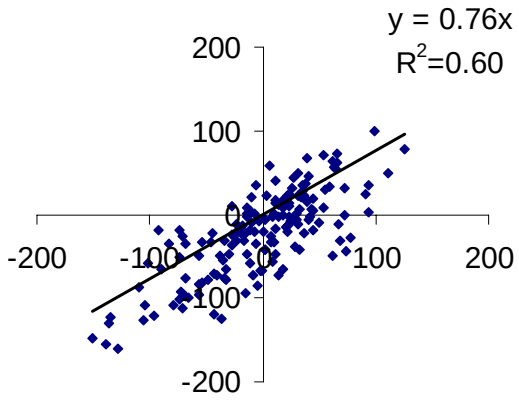
(a) Weekdays (Monday through Friday)

	Georgia (G)	Mountain (M)	Northeast (N)	Pacific (P)	Southeast (S)	midWest (W)
CO	1.18 ± 1.03	2.02 ± 0.24	0.57 ± 0.53	0.98 ± 1.20	1.35 ± 0.86	0.98 ± 7.97
NH3	0.44 ± 0.00	2.01 ± 0.01	0.46 ± 0.00	1.00 ± 0.00	0.51 ± 0.00	0.45 ± 0.01
NOXA	0.49 ± 0.01	0.56 ± 0.07	0.50 ± 0.01	1.79 ± 0.01	0.63 ± 0.01	0.50 ± 0.15
NOXP	0.46 ± 0.00	1.67 ± 0.03	0.85 ± 0.01	2.12 ± 0.00	0.44 ± 0.00	0.45 ± 0.38
PEC	0.56 ± 0.00	0.81 ± 0.00	0.45 ± 0.00	0.45 ± 0.00	0.81 ± 0.00	0.45 ± 0.00
PMFINE	0.45 ± 0.00	0.45 ± 0.00	0.45 ± 0.00	0.52 ± 0.00	0.46 ± 0.00	0.45 ± 0.00
POA	1.29 ± 0.00	1.38 ± 0.01	0.63 ± 0.00	1.36 ± 0.00	2.31 ± 0.00	1.32 ± 0.00
SO2	0.47 ± 0.00	0.48 ± 0.02	0.55 ± 0.00	1.43 ± 0.02	0.65 ± 0.00	0.88 ± 0.01
VOCA	0.37 ± 0.03	0.27 ± 0.00	0.56 ± 0.06	3.00 ± 1.32	3.20 ± 0.84	0.35 ± 17.2
VOCB	3.66 ± 0.00	0.37 ± 0.02	3.64 ± 0.01	0.49 ± 0.08	2.72 ± 0.04	3.70 ± 298.7

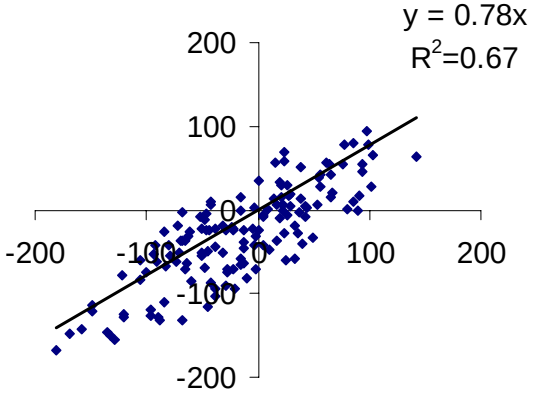
(b) Weekends (Saturday and Sunday)

	Georgia (G)	Mountain (M)	Northeast (N)	Pacific (P)	Southeast (S)	midWest (W)
CO	1.60 ± 0.73	2.19 ± 0.02	0.48 ± 0.26	1.00 ± 1.23	1.34 ± 0.72	0.64 ± 0.74
NH3	0.45 ± 0.00	0.46 ± 0.01	0.45 ± 0.01	1.53 ± 0.01	0.56 ± 0.00	0.45 ± 0.01
NOXA	0.48 ± 0.00	0.56 ± 0.04	0.48 ± 0.01	1.76 ± 0.01	0.63 ± 0.01	0.45 ± 0.00
NOXP	0.45 ± 0.00	0.45 ± 0.04	0.45 ± 0.00	2.20 ± 0.00	0.45 ± 0.00	0.48 ± 0.00
PEC	0.47 ± 0.00	2.13 ± 0.01	0.45 ± 0.00	0.61 ± 0.00	0.61 ± 0.00	0.45 ± 0.00
PMFINE	0.45 ± 0.00	0.44 ± 0.00	0.44 ± 0.00	1.28 ± 0.00	0.45 ± 0.00	0.45 ± 0.00
POA	0.57 ± 0.00	3.71 ± 0.00	0.27 ± 0.00	3.71 ± 0.00	2.55 ± 0.00	0.69 ± 0.00
SO2	0.46 ± 0.00	0.80 ± 0.02	0.71 ± 0.00	1.52 ± 0.04	0.78 ± 0.00	2.09 ± 0.01
VOCA	0.34 ± 0.10	0.27 ± 0.00	0.28 ± 0.01	3.70 ± 0.00	3.62 ± 0.05	0.42 ± 0.67
VOCB	3.67 ± 0.00	0.29 ± 0.00	3.63 ± 0.00	0.29 ± 0.01	2.75 ± 0.28	3.45 ± 0.00

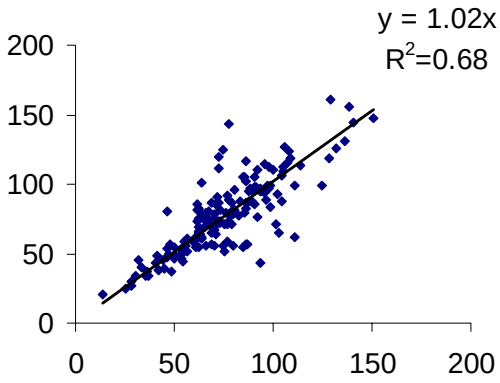
(a) MFB (July 2001)



(c) MFB (January 2002)



(b) MFE (July 2001)



(d) MFE (January 2002)

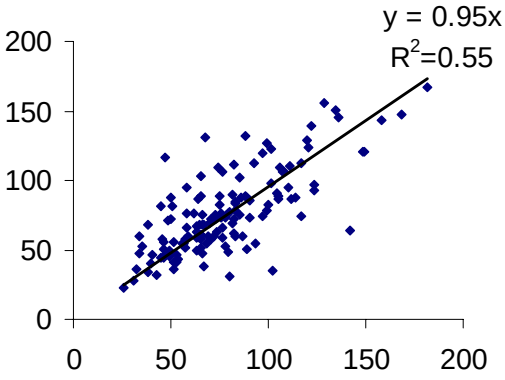


Figure 6-4. Comparison of the air quality model performance. The x-axis is the model performance with base emissions, and the y-axis is that with adjusted emissions. A slope less than 1.0 indicates that the model error (or bias) decreases, so the model performance improves. Equations for MFB and MFE are available in Table 3-1.

6.4. Conclusions

Emissions strengths biases were identified using FDDA. Regionally separate scaling factors were calculated to account for the regionally different biases. Further, a

separate scaling factor is calculated for weekdays and weekends to capture the day-of-week variation in the emission bias. This study considered the adjustment of the strength for 10 emission sources. Results showed that base emissions for CO and SO_2 sources were estimated reasonably well, while emissions for NH_3, NO_x, PEC and PMFINE, POA and VOC emissions require more significant revision. The results of emission adjustments increased the reliability of the air quality model, and thus enhanced the confidence of further applications of the model.

Appendix

Table 6A-1. Air quality model performance in the Pacific (P) region (July 2001)

			July 2001 (Base emissions)							July 2001 (Adjusted emissions)						
Species	No.	OBS	MODEL	MB	ME	NMB	NME	MFB	MFE	MODEL	MB	ME	NMB	NME	MFB	MFE
	OBS	conc.	conc	[μg /ppb]			[%]			conc	[μg /ppb]			[%]		
O_3	444	50.25	40.92	-9.3	13.5	-18.6	27.0	-23.4	31.6	39.57	-11.0	18.4	-21.8	36.3	-32.1	45.4
CO	778	452.53	327.87	-124.7	267.3	-27.5	59.1	-26.7	62.8	294.81	-166.6	262.9	-36.1	57.0	-36.2	63.7
NO_2	1040	11.48	12.25	0.8	7.2	6.7	62.3	-9.3	72.3	18.32	6.3	11.7	53.0	98.0	21.5	86.7
NO_x	901	16.64	15.93	-0.7	10.6	-4.3	63.7	-26.9	77.1	28.20	10.8	19.7	62.0	113.3	10.4	91.2
NO_y	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SO_2	222	2.44	1.17	-1.3	1.5	-52.1	61.3	-72.4	86.4	0.90	-1.6	1.8	-64.1	70.1	-92.5	101.7
HCHO	34	2.98	1.87	-1.1	1.6	-37.3	52.4	-26.1	54.3	2.01	-1.0	1.5	-32.4	50.1	-24.3	52.2
CCHO	34	1.24	0.91	-0.3	0.7	-26.4	58.3	-20.8	62.1	1.06	-0.2	0.7	-14.1	60.8	-16.8	65.2
Acetone	6	7.78	0.81	-7.0	7.0	-89.6	89.6	-151.0	151.0	0.90	-6.9	6.9	-88.4	88.4	-147.2	147.2
Ethylene	14	2.83	1.46	-1.4	1.8	-48.5	65.2	-79.1	92.3	0.95	-1.9	1.9	-66.5	68.7	-107.8	109.9
Isoprene	14	0.22	0.39	0.2	0.3	81.5	137.7	39.9	87.7	0.57	0.4	0.4	161.8	200.6	66.7	95.3
ALK1	14	4.96	1.00	-4.0	4.0	-79.9	79.9	-138.5	138.5	0.66	-4.3	4.3	-86.8	86.8	-155.6	155.6
ALK2	14	7.60	3.27	-4.3	5.3	-57.0	70.0	-73.6	98.2	4.74	-2.9	4.7	-37.6	62.2	-50.3	83.6
ALK3	14	3.06	4.34	1.3	3.0	41.8	96.6	-14.5	70.2	2.80	-0.3	2.2	-8.7	71.4	-47.9	78.2
ALK4	14	7.59	3.26	-4.3	5.3	-57.1	69.8	-96.6	107.9	2.11	-5.5	5.6	-72.2	74.3	-121.6	123.7
ALK5	14	3.21	2.97	-0.2	2.8	-7.4	85.9	-49.0	90.5	1.92	-1.3	2.3	-40.3	70.6	-78.6	97.5
ARO1	14	2.35	1.51	-0.8	1.9	-35.7	79.1	-56.1	107.4	0.97	-1.4	1.8	-58.5	74.9	-82.6	115.5
ARO2	2	1.64	0.78	-0.9	0.9	-52.8	56.9	-54.2	62.0	0.55	-1.1	1.1	-66.4	66.4	-81.4	81.4
OLE1	14	1.01	0.65	-0.4	0.7	-35.2	71.2	-73.5	97.5	0.43	-0.6	0.7	-57.8	70.9	-101.2	111.9
OLE2	14	0.19	0.58	0.4	0.4	209.0	236.2	64.7	85.8	0.44	0.3	0.3	135.5	167.2	55.2	79.9
TNMOC	77	56.18	23.84	-32.3	40.9	-57.6	72.9	-34.4	70.2	25.74	-31.2	39.8	-54.8	69.9	-31.8	71.5
SO_4^{2-} $PM_{2.5}$	132	1.38	1.49	0.1	0.9	7.6	64.5	9.9	60.2	1.31	-0.1	0.8	-5.3	58.9	0.2	58.8
NO_3^- $PM_{2.5}$	132	0.83	0.36	-0.5	0.6	-57.4	74.1	-91.9	111.0	0.74	-0.1	0.7	-11.1	88.0	-18.2	98.7
NH_4^+ $PM_{2.5}$	27	1.19	0.78	-0.4	0.8	-34.0	63.6	-6.8	59.5	0.89	-0.3	0.7	-25.1	61.5	1.4	58.8
OC $PM_{2.5}$	132	1.45	1.56	0.1	0.9	7.8	61.6	31.3	62.4	2.06	0.6	1.4	41.6	99.0	49.8	82.3
EC $PM_{2.5}$	132	0.25	0.51	0.3	0.3	108.0	122.7	50.4	69.1	0.25	0.0	0.1	2.8	49.4	-10.2	57.2
Crustal $PM_{2.5}$	19	2.40	3.01	0.6	2.7	25.5	112.8	65.2	98.2	2.89	0.5	2.6	20.3	110.3	62.3	97.3
Total $PM_{2.5}$	727	7.24	9.09	1.8	3.6	25.5	50.4	26.4	46.3	9.61	2.2	4.4	29.3	58.6	31.3	53.3

Table 6A-2. Air quality model performance in the Pacific (P) region (January 2002)

			January 2002 (Base emissions)							January 2002 (Adjusted emissions)						
Species	No. OBS	OBS conc.	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE
O_3	71	43.10	29.59	-13.5	13.9	-31.3	32.2	-45.7	46.5	28.74	-14.5	14.8	-33.5	34.2	-49.9	50.6
CO	1202	909.53	564.40	-345.1	539.0	-37.9	59.3	-50.6	75.9	560.41	-358.2	543.9	-39.0	59.2	-50.9	76.3
NO_2	1350	20.83	20.80	0.0	10.4	-0.2	49.9	-15.9	64.7	24.41	3.4	12.9	15.9	61.1	-1.3	68.5
NO_x	1212	49.05	32.21	-16.8	30.4	-34.3	61.9	-47.1	85.8	46.45	-3.8	36.6	-7.5	72.9	-23.7	88.1
NO_y	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SO_2	301	2.52	2.18	-0.3	1.9	-13.8	73.6	-38.0	83.4	2.81	0.2	2.2	7.8	84.3	-23.1	84.7
HCHO	41	2.51	0.85	-1.7	1.7	-66.1	69.8	-70.3	83.3	1.24	-1.3	1.7	-50.7	68.5	-51.6	83.8
CCHO	41	1.19	0.44	-0.7	0.8	-62.7	65.6	-82.9	90.2	0.61	-0.6	0.8	-49.1	63.0	-67.4	85.8
Acetone	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Ethylene	19	7.78	2.71	-5.1	5.4	-65.2	68.9	-101.0	104.7	4.29	-3.5	5.5	-44.9	71.0	-74.6	91.1
Isoprene	19	0.21	0.13	-0.1	0.2	-38.8	86.7	-61.8	105.0	0.20	0.0	0.2	-7.0	82.1	-25.6	86.7
ALK1	19	13.30	1.41	-11.9	11.9	-89.4	89.4	-158.4	158.4	2.11	-11.2	11.2	-84.1	84.1	-143.4	143.4
ALK2	13	14.96	3.73	-11.2	11.5	-75.1	76.7	-105.3	110.3	5.32	-9.6	10.2	-64.4	68.4	-83.0	95.0
ALK3	19	13.50	7.92	-5.6	6.5	-41.3	48.1	-47.8	57.6	13.07	-0.4	5.8	-3.2	43.1	-12.2	52.0
ALK4	17	16.55	6.64	-9.9	9.9	-59.9	59.9	-90.8	90.8	10.46	-6.1	9.2	-36.8	55.4	-61.0	73.4
ALK5	17	6.17	6.39	0.2	3.2	3.7	51.8	-15.7	51.7	9.78	3.6	5.6	58.6	90.2	15.1	56.0
ARO1	17	5.68	3.24	-2.4	2.5	-43.1	44.2	-62.8	65.4	5.00	-0.7	3.1	-11.9	55.0	-31.9	60.8
ARO2	2	9.09	4.00	-5.1	5.1	-56.0	56.0	-79.5	79.5	6.08	-3.0	3.5	-33.2	38.1	-43.0	48.6
OLE1	19	2.74	1.46	-1.3	1.3	-46.8	47.1	-69.9	70.4	2.48	-0.3	1.8	-9.5	64.9	-36.3	72.3
OLE2	19	0.78	1.24	0.5	0.8	59.3	98.5	23.9	75.0	2.26	1.5	1.7	190.4	214.4	58.6	88.1
TNMOC	139	118.13	31.47	-86.7	90.3	-73.4	76.4	-94.0	105.3	60.60	-54.1	85.2	-47.2	74.3	-52.9	88.5
SO_4^{2-} $PM_{2.5}$	198	0.56	0.58	0.0	0.3	4.8	58.7	26.5	72.5	0.61	0.1	0.3	9.1	59.3	29.4	72.5
NO_3^- $PM_{2.5}$	174	2.57	2.93	0.4	2.4	13.8	91.9	38.9	105.8	3.31	0.7	2.5	28.2	96.1	51.8	108.9
NH_4^+ $PM_{2.5}$	44	2.50	1.92	-0.6	1.4	-23.4	54.2	20.3	72.6	2.17	-0.3	1.3	-13.2	52.2	29.4	72.7
OC $PM_{2.5}$	193	2.53	1.71	-0.8	2.0	-32.4	77.5	19.7	88.7	2.06	-0.5	2.0	-18.9	77.8	33.3	87.9
EC $PM_{2.5}$	198	0.42	0.64	0.2	0.4	52.7	88.8	16.7	80.5	0.32	-0.1	0.3	-22.9	60.1	-37.3	77.4
Crustal $PM_{2.5}$	40	5.27	4.93	-0.3	3.5	-6.4	67.0	13.7	59.2	4.86	-0.4	3.5	-7.7	66.0	14.1	59.2
Total $PM_{2.5}$	1029	14.53	13.50	-1.0	7.7	-7.1	53.3	-3.7	63.4	14.54	-0.2	8.0	-1.5	54.0	2.4	62.5

Table 6A-3. Air quality model performance in the Mountain (M) region (July 2001)

			July 2001 (Base emissions)							July 2001 (Adjusted emissions)						
Species	No. OBS	OBS conc.	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE
O_3	501	46.40	45.59	-0.8	6.3	-1.8	13.6	-2.7	13.8	44.25	-2.7	9.5	-5.7	20.3	-8.2	21.0
CO	472	565.50	404.70	-160.8	318.0	-28.4	56.2	-37.2	65.3	361.28	-215.8	326.1	-37.4	56.5	-49.3	69.4
NO_2	308	11.94	11.39	-0.6	7.5	-4.6	62.7	-11.7	71.2	6.70	-5.5	7.3	-45.1	60.1	-58.0	85.5
NO_x	145	16.50	9.63	-6.9	14.0	-41.6	85.0	-56.4	108.4	4.83	-11.3	12.9	-70.1	79.9	-95.8	118.3
NO_y	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SO_2	152	6.03	1.15	-4.9	4.9	-80.9	81.9	-101.2	104.5	2.23	-5.3	6.1	-70.2	80.8	-59.2	87.4
HCHO	7	5.75	48.70	42.9	43.4	746.9	754.5	110.9	125.0	14.62	8.9	10.3	154.3	179.4	49.1	98.6
CCHO	7	1.36	19.15	17.8	17.9	1310.9	1316.1	125.2	132.0	6.08	4.7	5.2	347.7	382.2	78.2	125.8
Acetone	7	2.92	10.45	7.5	8.3	257.3	283.8	67.5	111.0	3.26	0.3	1.6	11.4	56.3	-12.6	62.0
Ethylene	6	2.44	17.33	14.9	16.1	610.3	661.6	-15.7	77.2	4.74	2.3	5.2	94.3	214.5	-93.5	143.1
Isoprene	6	0.35	1.25	0.9	1.0	255.6	274.3	78.2	90.9	0.32	0.0	0.3	-9.1	95.1	-27.3	85.8
ALK1	6	2.34	2.88	0.5	2.1	23.0	89.1	-42.8	72.1	0.87	-1.5	1.5	-63.0	63.0	-119.9	119.9
ALK2	6	3.93	11.91	8.0	8.4	202.7	212.2	72.1	77.7	4.02	0.1	3.0	2.3	76.2	0.0	73.4
ALK3	6	2.53	7.73	5.2	5.2	205.3	206.9	68.9	87.3	2.17	-0.4	1.1	-14.2	43.7	-31.0	57.1
ALK4	6	8.26	9.80	1.5	6.3	18.6	75.8	-37.2	74.3	2.70	-5.6	5.6	-67.3	67.3	-124.7	124.7
ALK5	6	2.37	11.02	8.6	9.2	364.3	387.9	14.0	85.9	2.98	0.6	3.2	25.5	133.9	-72.6	116.9
ARO1	6	2.00	12.49	10.5	11.2	525.1	561.3	18.3	95.4	3.39	1.4	3.4	69.8	171.9	-65.5	114.0
ARO2	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
OLE1	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
OLE2	6	0.55	1.68	1.1	1.3	203.5	239.7	61.8	91.8	0.44	-0.1	0.4	-20.9	76.6	-49.1	76.2
TNMOC	6	47.79	139.85	92.1	99.7	192.7	208.7	32.3	46.5	45.23	-2.6	42.7	-5.4	89.3	-42.9	80.5
SO_4^{2-} $PM_{2.5}$	173	1.11	1.14	0.0	0.5	2.8	42.0	14.1	41.5	1.11	0.0	0.5	0.6	49.3	11.2	48.1
NO_3^- $PM_{2.5}$	173	0.20	0.05	-0.2	0.2	-76.4	84.6	-134.8	140.4	0.07	-0.1	0.2	-64.5	93.0	-122.9	144.4
NH_4^+ $PM_{2.5}$	21	0.42	0.49	0.1	0.3	18.2	60.8	40.5	62.4	0.54	0.1	0.3	29.7	70.4	45.7	69.4
OC $PM_{2.5}$	175	1.07	1.19	0.1	0.7	10.8	64.1	21.3	56.6	0.88	-0.2	0.6	-17.7	58.6	-19.9	60.9
EC $PM_{2.5}$	176	0.17	0.23	0.1	0.1	36.8	87.4	11.8	61.3	0.12	-0.1	0.1	-29.4	65.2	-48.0	73.6
Crustal $PM_{2.5}$	16	4.06	2.52	-1.5	3.1	-37.9	75.6	11.5	78.3	2.74	-1.3	3.1	-32.6	75.4	16.4	76.9
Total $PM_{2.5}$	607	5.41	6.68	1.3	2.9	23.4	53.7	18.4	44.5	6.18	0.7	2.9	11.9	51.7	7.2	45.5

Table 6A-4. Air quality model performance in the Mountain (M) region (January 2002)

			January 2002 (Base emissions)							January 2002 (Adjusted emissions)						
Species	No. OBS	OBS conc.	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE	MODEL conc	MB [μg /ppb]	ME	NMB	NME	MFB [%]	MFE
O_3	104	44.41	36.43	-8.0	9.3	-18.0	21.0	-23.0	26.0	37.09	-7.4	8.5	-16.7	19.0	-19.8	22.2
CO	920	1074.17	440.47	-633.7	667.8	-59.0	62.2	-79.3	85.0	647.55	-440.6	627.3	-40.5	57.7	-54.2	75.0
NO_2	379	24.18	16.42	-7.8	11.1	-32.1	46.1	-39.0	65.1	8.55	-15.5	16.3	-64.5	67.9	-94.1	103.1
NO_x	124	63.63	16.52	-47.1	52.4	-74.0	82.4	-89.0	121.9	7.29	-57.4	58.6	-88.7	90.5	-127.8	139.1
NO_y	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SO_2	215	4.97	1.50	-3.5	3.7	-69.8	74.4	-83.6	92.6	1.08	-4.1	4.1	-79.0	80.4	-109.6	112.6
HCHO	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CCHO	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Acetone	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Ethylene	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Isoprene	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ALK1	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ALK2	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ALK3	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ALK4	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ALK5	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ARO1	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ARO2	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
OLE1	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
OLE2	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
TNMOC	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SO_4^{2-} $PM_{2.5}$	252	0.43	0.58	0.1	0.3	32.8	66.2	20.5	63.1	0.53	0.1	0.3	21.2	57.6	14.8	59.2
NO_3^- $PM_{2.5}$	241	0.80	1.41	0.6	1.2	75.8	145.7	64.1	110.8	1.24	0.4	1.1	54.5	137.3	54.8	110.3
NH_4^+ $PM_{2.5}$	40	0.80	0.74	-0.1	0.5	-7.5	61.0	4.5	73.0	0.67	-0.1	0.5	-15.8	62.9	-9.0	75.0
OC $PM_{2.5}$	251	1.20	0.82	-0.4	1.0	-31.1	80.2	16.7	83.1	0.74	-0.5	0.9	-38.5	76.6	5.6	78.6
EC $PM_{2.5}$	252	0.28	0.24	0.0	0.2	-12.7	58.5	-12.9	66.2	0.17	-0.1	0.2	-39.6	59.6	-46.2	74.9
Crustal $PM_{2.5}$	25	1.40	2.82	1.4	2.0	101.1	144.6	61.1	107.0	2.66	1.3	1.9	89.7	134.8	57.4	105.8
Total $PM_{2.5}$	714	7.09	6.95	-0.1	4.5	-1.9	63.2	8.6	70.2	6.37	-0.9	4.4	-12.7	59.8	0.1	68.6

Table 6A-5. Air quality model performance in the midWestern (W) region (July 2001)

			July 2001 (Base emissions)							July 2001 (Adjusted emissions)						
Species	No. OBS	OBS conc.	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE
O_3	1330	49.60	63.67	14.1	16.4	28.4	33.2	23.0	27.9	58.45	9.0	14.5	18.3	29.3	13.7	26.5
CO	644	569.43	424.80	-144.6	286.1	-25.4	50.2	-22.8	55.7	433.52	-127.3	293.3	-22.7	52.3	-22.0	58.3
NO_2	502	14.88	15.19	0.3	7.1	2.0	47.9	-14.6	53.7	17.30	2.4	8.2	16.4	55.0	-2.5	54.0
NO_x	289	20.73	19.45	-1.3	10.2	-6.2	49.2	-13.4	54.9	26.62	4.7	15.4	21.4	70.0	0.6	58.9
NO_y	19	11.19	17.83	6.6	11.9	59.3	106.7	23.3	73.1	19.28	8.4	13.3	77.5	122.2	24.1	71.8
SO_2	1080	5.85	4.51	-1.3	4.2	-23.0	71.4	-32.7	80.6	2.33	-3.1	3.8	-57.4	70.1	-78.2	96.3
HCHO	73	3.12	3.68	0.6	1.8	17.8	56.5	23.1	54.5	2.72	-0.3	1.4	-10.8	44.8	-2.3	50.0
CCHO	63	0.95	1.61	0.7	0.9	69.0	92.6	59.4	71.2	1.20	0.2	0.7	26.0	70.9	33.4	64.2
Acetone	34	0.59	1.21	0.6	0.8	103.5	139.2	72.2	96.4	0.77	0.1	0.6	22.8	90.7	31.4	89.0
Ethylene	73	1.60	0.96	-0.6	0.9	-40.3	58.2	-65.7	84.6	0.61	-1.0	1.1	-62.5	65.9	-99.5	104.9
Isoprene	67	0.68	0.40	-0.3	0.6	-40.8	91.0	1.0	104.1	0.30	-0.4	0.6	-55.1	89.7	-20.8	106.2
ALK1	84	4.58	1.23	-3.3	3.3	-73.1	73.1	-105.5	105.5	0.81	-3.6	3.6	-81.6	81.6	-127.0	127.0
ALK2	84	3.60	3.52	-0.1	2.3	-2.2	64.5	31.6	48.5	2.18	-1.3	1.8	-37.6	52.7	-11.3	37.4
ALK3	84	3.15	4.34	1.2	4.4	37.8	138.6	43.4	80.3	2.73	-0.3	3.3	-9.1	111.2	5.9	81.5
ALK4	84	4.50	3.43	-1.1	3.6	-23.7	80.5	-40.6	90.6	2.22	-2.2	3.3	-49.6	75.6	-72.0	105.4
ALK5	73	1.70	2.48	0.8	1.6	45.9	92.3	6.7	67.1	1.56	-0.1	1.2	-7.4	71.4	-33.3	80.9
ARO1	74	1.86	2.42	0.6	2.0	30.0	105.4	7.0	77.8	1.61	-0.2	1.6	-12.4	85.8	-26.7	89.1
ARO2	53	2.72	1.41	-1.3	2.3	-48.2	85.9	-33.0	88.7	0.90	-1.8	2.2	-66.1	82.2	-65.2	97.2
OLE1	50	0.88	0.51	-0.4	0.5	-42.6	56.4	-68.6	85.8	0.34	-0.5	0.6	-61.8	66.3	-97.2	105.0
OLE2	51	0.27	0.16	-0.1	0.1	-38.2	51.7	-43.7	69.1	0.12	-0.1	0.2	-54.1	63.8	-70.4	86.3
TNMOC	84	54.45	43.20	-11.3	48.4	-20.7	88.8	23.4	79.0	26.60	-24.5	40.3	-48.0	78.8	-11.1	71.1
SO_4^{2-} $PM_{2.5}$	104	5.60	8.02	2.4	3.4	43.2	61.2	30.5	50.3	5.53	-0.1	2.3	-1.3	41.3	0.3	46.2
NO_3^- $PM_{2.5}$	102	0.69	0.77	0.1	0.5	11.7	71.3	-15.8	73.5	0.73	0.0	0.5	6.8	78.0	-22.7	81.1
NH_4^+ $PM_{2.5}$	58	2.12	3.28	1.2	1.5	54.7	72.7	55.2	66.0	2.38	0.3	1.1	11.9	49.8	29.1	55.2
OC $PM_{2.5}$	105	3.66	1.41	-2.2	2.4	-61.4	65.6	-74.3	86.2	3.02	-0.6	2.0	-17.5	53.8	-18.0	57.0
EC $PM_{2.5}$	105	0.48	0.83	0.4	0.4	74.1	91.4	36.9	62.4	0.43	0.0	0.2	-9.9	41.7	-21.8	54.5
Crustal $PM_{2.5}$	54	5.00	3.81	-1.2	3.3	-23.9	65.2	-18.4	68.5	3.55	-1.5	3.3	-29.0	65.0	-24.7	70.0
Total $PM_{2.5}$	1358	20.76	21.90	1.1	7.2	5.5	34.8	1.9	36.3	19.04	-1.0	6.9	-5.0	34.2	-6.2	37.4

Table 6A-6. Air quality model performance in the midWestern (W) region (January 2002)

			January 2002 (Base emissions)							January 2002 (Adjusted emissions)						
Species	No.	OBS	MODEL	MB	ME	NMB	NME	MFB	MFE	MODEL	MB	ME	NMB	NME	MFB	MFE
	OBS	conc.	conc	[μg /ppb]			[%]			conc	[μg /ppb]			[%]		
O_3	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CO	811	626.68	501.55	-125.1	325.2	-20.0	51.9	-21.8	56.1	486.48	-125.0	309.3	-20.4	50.6	-21.6	54.5
NO_2	588	15.38	14.27	-1.1	4.7	-7.2	30.8	-12.2	38.4	8.63	-6.4	6.9	-42.6	46.1	-63.0	67.8
NO_x	300	32.19	28.25	-3.9	17.2	-12.2	53.5	-29.0	58.4	12.89	-18.2	19.4	-58.5	62.4	-90.4	95.0
NO_y	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SO_2	1636	5.65	4.82	-0.8	3.6	-14.7	64.4	-13.2	66.0	4.10	-1.6	3.6	-27.7	63.0	-25.0	67.5
HCHO	52	1.32	1.08	-0.2	0.8	-17.9	64.4	-22.8	75.2	0.61	-0.8	0.8	-55.4	61.4	-70.9	82.8
CCHO	42	0.84	0.48	-0.4	0.4	-43.0	46.3	-42.3	49.9	0.29	-0.6	0.6	-65.9	66.4	-86.8	88.1
Acetone	66	1.71	0.43	-1.3	1.3	-75.2	76.4	-95.9	99.4	0.27	-1.4	1.4	-84.3	84.3	-126.6	126.6
Ethylene	6	1.64	1.65	0.0	0.8	0.4	45.9	-10.6	44.7	0.69	-0.9	0.9	-57.7	57.7	-81.2	81.2
Isoprene	3	0.02	0.01	0.0	0.0	-48.8	48.8	-67.7	67.7	0.00	0.0	0.0	-81.2	81.2	-131.3	131.3
ALK1	6	5.16	1.18	-4.0	4.0	-77.2	77.2	-128.9	128.9	0.64	-4.5	4.5	-87.6	87.6	-155.5	155.5
ALK2	3	5.63	2.12	-3.5	3.5	-62.3	62.3	-88.5	88.5	1.08	-4.5	4.5	-80.8	80.8	-131.7	131.7
ALK3	6	2.74	4.02	1.3	1.3	46.5	46.5	39.8	39.8	1.68	-1.1	1.1	-38.6	38.6	-40.0	40.0
ALK4	3	4.05	5.10	1.0	1.8	25.8	45.6	33.2	45.7	1.93	-2.1	2.1	-52.4	52.4	-57.8	57.8
ALK5	6	1.18	2.69	1.5	1.5	127.6	127.6	80.4	80.4	1.08	-0.1	0.4	-8.7	30.0	0.3	30.9
ARO1	6	0.86	2.61	1.8	1.8	204.4	204.4	102.1	102.1	1.06	0.2	0.3	23.1	36.0	27.6	34.7
ARO2	3	1.07	2.37	1.3	1.3	121.8	121.8	88.9	88.9	0.84	-0.2	0.6	-21.7	52.5	-0.8	50.6
OLE1	6	1.29	0.85	-0.4	0.5	-34.2	36.2	-44.9	47.2	0.33	-1.0	1.0	-74.1	74.1	-116.1	116.1
OLE2	2	0.15	0.86	0.7	0.7	493.1	493.1	141.8	141.8	0.28	0.1	0.1	94.9	94.9	63.7	63.7
TNMOC	30	107.01	18.63	-88.4	88.9	-82.6	83.1	-132.5	134.7	12.76	-94.3	94.3	-88.1	88.1	-150.4	150.4
SO_4^{2-} $PM_{2.5}$	172	2.18	1.28	-0.9	1.2	-41.4	53.0	-50.5	67.1	1.22	-1.0	1.2	-44.2	53.4	-54.1	68.2
NO_3^- $PM_{2.5}$	171	2.93	4.64	1.7	2.7	58.8	91.1	36.6	78.3	2.84	-0.1	1.8	-3.0	62.2	-2.6	73.5
NH_4^+ $PM_{2.5}$	111	1.70	2.14	0.4	1.0	25.8	59.8	19.6	58.5	1.51	-0.2	0.9	-11.2	52.6	-10.6	60.2
OC $PM_{2.5}$	174	2.43	1.36	-1.1	1.4	-44.2	57.4	-50.5	76.2	1.90	-0.5	1.3	-22.0	53.4	-23.2	65.6
EC $PM_{2.5}$	174	0.35	0.70	0.3	0.4	98.7	117.3	33.6	70.5	0.36	0.0	0.2	3.0	52.3	-18.9	57.6
Crustal $PM_{2.5}$	86	1.60	4.27	2.7	3.0	167.0	189.4	85.9	101.5	3.88	2.3	2.7	142.4	168.5	79.7	97.9
Total $PM_{2.5}$	1733	12.03	15.90	3.9	6.2	32.2	51.7	22.1	46.3	13.12	1.2	5.2	10.1	43.6	5.7	44.1

Table 6A-7. Air quality model performance in the Northeastern (N) US (July 2001)

			July 2001 (Base emissions)							July 2001 (Adjusted emissions)						
Species	No. OBS	OBS conc.	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE
O_3	615	51.70	62.90	11.2	14.7	21.7	28.3	17.8	25.5	62.26	10.6	14.1	20.5	27.3	17.0	24.7
CO	625	544.89	401.19	-143.7	268.9	-26.4	49.4	-27.5	55.6	411.82	-126.4	269.9	-23.5	50.1	-24.7	55.6
NO_2	682	15.36	15.04	-0.3	8.3	-2.1	53.8	-16.0	62.2	12.24	-2.6	7.6	-17.5	51.0	-31.5	65.6
NO_x	579	20.59	17.32	-3.3	11.3	-15.9	54.8	-30.3	67.4	14.04	-6.3	11.3	-30.9	55.4	-46.7	73.8
NO_y	50	9.46	9.34	-0.1	3.6	-1.2	38.1	13.4	35.7	7.99	-1.5	3.5	-15.5	36.8	-2.6	34.4
SO_2	893	5.56	6.51	1.0	4.2	17.1	75.6	18.1	71.3	4.91	-0.4	3.7	-8.0	69.9	-1.6	70.6
HCHO	25	4.59	3.70	-0.9	2.0	-19.3	42.8	-7.4	49.8	3.17	-1.4	2.0	-30.8	44.5	-20.5	54.3
CCHO	28	2.09	1.90	-0.2	1.7	-9.2	81.8	39.3	70.6	1.48	-0.6	1.4	-29.4	68.4	16.0	55.6
Acetone	30	4.33	1.29	-3.0	3.3	-70.3	76.6	-68.7	97.4	1.17	-3.2	3.4	-73.1	77.5	-76.3	99.1
Ethylene	111	0.94	1.29	0.4	0.8	37.6	80.8	12.2	61.2	1.28	0.3	0.7	37.2	79.3	13.8	61.9
Isoprene	117	0.52	0.89	0.4	0.7	72.9	127.3	11.1	72.5	0.93	0.4	0.7	88.0	140.8	14.3	73.4
ALK1	130	2.16	0.63	-1.5	1.5	-70.9	71.3	-104.7	105.4	0.60	-1.5	1.5	-71.9	72.3	-108.3	108.8
ALK2	128	2.03	2.45	0.4	1.3	20.7	64.2	31.7	54.4	2.16	0.2	1.2	9.8	58.6	21.6	51.5
ALK3	123	0.92	4.54	3.6	3.9	396.1	424.7	99.4	113.9	4.38	3.5	3.7	394.4	421.0	99.4	113.0
ALK4	124	1.81	3.32	1.5	2.5	83.7	138.7	21.3	76.2	3.26	1.5	2.5	86.9	142.1	21.9	79.0
ALK5	109	0.78	2.28	1.5	1.8	191.5	225.1	61.8	93.7	2.29	1.5	1.8	196.1	229.1	63.6	95.0
ARO1	128	0.82	1.91	1.1	1.3	134.5	161.3	35.9	77.6	1.88	1.1	1.3	133.0	160.5	35.5	78.5
ARO2	129	0.52	1.21	0.7	0.9	132.2	178.3	37.2	87.8	1.20	0.7	0.9	128.7	174.5	37.3	87.7
OLE1	103	0.49	0.49	0.0	0.5	-0.5	110.6	-17.7	88.8	0.48	0.0	0.5	3.0	110.1	-13.1	87.5
OLE2	109	0.58	0.36	-0.2	0.7	-37.8	116.1	-10.5	94.1	0.36	-0.2	0.7	-39.3	114.5	-9.4	95.2
TNMOC	162	25.94	39.58	13.6	28.6	52.6	110.1	63.7	86.0	36.13	10.8	26.3	42.4	103.8	56.7	82.7
SO_4^{2-} $PM_{2.5}$	147	7.91	9.64	1.7	2.8	21.8	36.0	34.1	45.1	7.78	0.2	2.3	2.9	30.1	20.9	39.3
NO_3^- $PM_{2.5}$	147	0.53	0.77	0.2	0.6	46.3	111.1	-35.7	100.1	0.64	0.1	0.6	25.0	112.4	-55.1	110.7
NH_4^+ $PM_{2.5}$	85	2.56	3.11	0.6	1.1	21.8	44.5	36.0	50.6	2.37	0.0	1.0	0.8	42.3	20.8	48.5
OC $PM_{2.5}$	151	3.70	1.73	-2.0	2.1	-53.2	56.8	-68.2	74.5	2.69	-1.0	1.8	-27.0	47.9	-35.7	55.6
EC $PM_{2.5}$	151	0.49	0.91	0.4	0.5	85.1	106.9	34.9	63.7	0.85	0.4	0.5	73.2	98.6	28.3	60.5
Crustal $PM_{2.5}$	82	7.28	3.74	-3.5	4.9	-48.7	66.8	-45.3	80.0	3.63	-3.1	4.4	-45.8	66.2	-33.7	78.5
Total $PM_{2.5}$	848	16.84	20.53	3.7	7.6	21.9	45.1	21.7	40.6	19.83	3.6	8.0	22.3	49.6	22.1	43.5

Table 6A-8. Air quality model performance in the Northeastern (N) US (January 2002)

			January 2002 (Base emissions)							January 2002 (Adjusted emissions)						
Species	No. OBS	OBS conc.	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE
O_3	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CO	805	716.64	742.53	25.9	392.3	3.6	54.7	-4.6	49.3	574.50	-119.2	333.3	-17.2	48.0	-22.3	49.6
NO_2	870	22.64	21.77	-0.9	6.9	-3.9	30.5	-3.8	35.5	14.79	-6.9	8.7	-31.7	40.0	-42.0	53.1
NO_x	718	44.94	39.60	-5.3	25.6	-11.9	56.9	-15.6	58.0	20.42	-21.6	24.6	-51.4	58.6	-66.1	76.7
NO_y	51	28.89	18.63	-10.3	18.3	-35.5	63.2	4.8	57.2	9.77	-17.3	18.6	-63.9	68.5	-42.1	57.3
SO_2	1333	8.36	10.76	2.4	5.3	28.7	63.9	21.6	56.4	8.14	0.0	4.4	0.5	54.4	-1.5	54.9
HCHO	31	2.30	1.30	-1.0	1.4	-43.6	60.4	-63.6	79.3	1.29	-1.0	1.4	-43.9	58.8	-64.1	75.9
CCHO	31	0.94	0.65	-0.3	0.3	-30.5	31.4	-39.0	40.3	0.66	-0.3	0.3	-29.3	32.3	-43.6	46.6
Acetone	19	2.21	0.40	-1.8	1.8	-81.7	81.7	-120.5	120.5	0.39	-1.8	1.8	-82.3	82.3	-124.1	124.1
Ethylene	21	3.83	1.99	-1.8	2.1	-48.0	54.0	-47.4	61.7	1.61	-2.2	2.3	-57.8	60.9	-68.9	76.0
Isoprene	18	0.04	0.02	0.0	0.0	-59.4	64.7	-65.1	81.9	0.02	0.0	0.0	-61.8	68.7	-70.3	89.2
ALK1	21	6.86	1.16	-5.7	5.7	-83.0	83.0	-135.8	135.8	1.00	-5.9	5.9	-85.4	85.4	-145.5	145.5
ALK2	21	7.90	1.63	-6.3	6.3	-79.4	79.4	-120.0	120.0	1.47	-6.4	6.4	-81.5	81.5	-128.8	128.8
ALK3	21	4.04	5.21	1.2	1.8	28.7	45.3	38.0	42.7	4.32	0.3	1.3	6.7	32.6	13.8	31.8
ALK4	21	3.86	4.64	0.8	0.9	20.4	23.9	28.8	30.6	3.82	0.0	0.9	-1.0	22.1	4.2	28.1
ALK5	21	1.29	3.27	2.0	2.0	154.1	154.1	98.9	98.9	2.66	1.4	1.4	107.0	107.2	78.7	78.8
ARO1	21	1.62	3.10	1.5	1.5	91.2	91.2	66.0	66.0	2.54	0.9	1.0	56.4	61.2	41.6	47.4
ARO2	21	1.10	2.24	1.1	1.1	103.8	103.8	76.7	76.7	1.83	0.7	0.8	66.7	72.4	54.1	58.9
OLE1	21	1.90	0.86	-1.0	1.1	-54.9	57.0	-59.8	65.5	0.66	-1.2	1.3	-65.3	66.3	-85.2	88.3
OLE2	21	0.39	0.60	0.2	0.2	56.6	57.1	53.5	54.0	0.37	0.0	0.1	-5.2	31.5	6.0	43.2
TNMOC	66	58.22	65.38	7.2	31.5	12.3	54.1	1.3	48.9	47.33	-10.1	25.2	-17.6	44.0	-23.9	47.3
SO_4^{2-} $PM_{2.5}$	194	2.45	2.12	-0.3	1.1	-13.4	46.0	-22.7	48.1	2.01	-0.4	1.1	-18.2	42.9	-23.6	47.8
NO_3^- $PM_{2.5}$	194	2.33	7.18	4.8	4.9	207.8	209.6	93.6	97.2	4.80	2.5	2.7	104.7	117.1	55.5	73.9
NH_4^+ $PM_{2.5}$	112	1.73	3.26	1.5	1.6	88.5	93.8	55.5	63.1	2.41	0.7	1.0	40.3	59.2	27.0	49.5
OC $PM_{2.5}$	193	2.66	2.24	-0.4	1.5	-16.0	58.2	-31.3	64.1	2.42	-0.2	1.5	-8.9	57.3	-19.6	59.9
EC $PM_{2.5}$	193	0.55	1.38	0.8	0.9	150.6	168.5	42.4	72.3	0.90	0.4	0.5	65.7	99.3	3.9	60.2
Crustal $PM_{2.5}$	84	1.77	7.17	5.4	5.6	305.1	314.4	97.9	108.4	6.73	5.0	5.2	286.4	297.7	94.8	106.2
Total $PM_{2.5}$	1124	12.43	24.70	12.3	14.0	98.8	112.3	54.2	68.8	20.92	8.6	11.1	69.8	89.7	38.7	60.2

Table 6A-9. Air quality model performance in the Southeastern (S) US (July 2001)

			July 2001 (Base emissions)							July 2001 (Adjusted emissions)						
Species	No. OBS	OBS conc.	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE
O_3	1407	47.66	61.15	13.5	15.9	28.3	33.4	22.8	28.5	54.98	7.3	15.4	15.2	32.2	9.2	29.9
CO	1004	523.24	296.49	-226.7	270.4	-43.3	51.7	-44.3	60.8	391.41	-129.0	255.5	-24.8	49.1	-22.4	54.4
NO_2	1039	10.32	10.94	0.6	6.7	6.0	64.5	-5.8	66.2	17.16	6.9	10.7	67.2	103.8	35.2	75.9
NO_x	536	12.12	13.12	1.0	7.4	8.2	61.3	5.3	61.6	25.23	13.2	15.9	109.8	132.6	58.7	81.2
NO_y	231	9.34	9.84	0.5	5.5	5.4	59.2	11.1	62.4	15.38	6.0	9.0	64.5	95.9	40.7	74.8
SO_2	1050	5.35	5.39	0.0	3.7	0.6	69.5	1.5	65.8	3.34	-2.0	3.4	-37.1	64.6	-41.7	73.1
HCHO	30	4.54	3.65	-0.9	2.4	-19.6	52.8	-6.5	48.1	2.53	-2.0	2.5	-44.2	55.7	-35.8	57.2
CCHO	29	0.86	1.57	0.7	0.9	83.2	110.0	40.6	70.6	0.83	0.0	0.5	-2.8	63.5	-5.6	66.7
Acetone	27	0.63	0.98	0.4	0.6	56.1	91.1	44.9	70.5	0.70	0.1	0.5	11.6	80.8	18.2	75.2
Ethylene	186	2.70	0.80	-1.9	2.2	-70.2	80.5	-54.1	89.8	1.07	-1.7	2.2	-61.4	81.0	-33.1	86.6
Isoprene	166	0.54	0.68	0.1	0.4	25.4	80.8	29.1	91.6	0.28	-0.2	0.4	-46.7	70.7	-38.5	98.0
ALK1	186	6.40	0.70	-5.7	5.7	-89.1	89.1	-136.0	136.4	0.78	-5.5	5.5	-87.7	87.8	-130.6	131.4
ALK2	186	6.37	2.93	-3.4	4.3	-54.0	68.1	-34.0	71.4	1.60	-4.6	4.8	-74.3	77.5	-76.6	90.5
ALK3	186	4.47	2.74	-1.7	4.3	-38.6	95.8	3.3	98.5	3.53	-0.7	4.4	-16.1	104.4	22.9	99.4
ALK4	177	4.39	2.22	-2.2	3.6	-49.4	82.6	-11.7	95.3	2.91	-1.4	3.7	-32.6	85.9	4.2	93.7
ALK5	185	1.01	1.44	0.4	1.0	42.4	102.4	54.4	90.2	1.91	0.9	1.3	88.0	131.8	70.0	95.0
ARO1	184	1.24	0.98	-0.3	0.9	-20.9	70.4	-13.5	79.1	1.29	0.1	1.0	5.1	80.9	8.3	80.4
ARO2	71	0.69	0.79	0.1	0.6	14.6	86.3	26.9	83.6	1.05	0.4	0.7	51.0	106.7	43.5	87.9
OLE1	178	12.74	0.37	-12.4	12.4	-97.1	97.6	-89.2	102.3	0.52	-11.9	12.1	-95.8	97.2	-64.8	92.9
OLE2	159	0.29	0.36	0.1	0.3	22.0	108.7	65.6	105.0	0.49	0.2	0.4	67.6	143.5	73.1	112.0
TNMOC	144	33.81	35.26	1.5	28.8	4.3	85.3	32.5	76.5	25.63	-7.1	24.7	-21.6	75.6	6.6	70.9
SO_4^{2-} $PM_{2.5}$	269	6.35	8.85	2.5	2.9	39.2	46.0	30.5	42.0	5.92	-0.3	2.0	-5.1	32.5	-6.6	38.1
NO_3^- $PM_{2.5}$	193	0.33	0.20	-0.1	0.3	-39.5	80.6	-109.9	128.1	0.26	-0.1	0.3	-20.3	85.1	-87.8	118.2
NH_4^+ $PM_{2.5}$	197	1.82	2.32	0.5	0.8	27.4	45.4	28.2	46.1	1.60	-0.2	0.7	-9.9	39.4	-6.6	47.1
OC $PM_{2.5}$	274	3.24	1.35	-1.9	1.9	-58.3	58.8	-83.4	85.1	2.45	-0.8	1.4	-24.6	42.2	-35.7	54.4
EC $PM_{2.5}$	273	0.53	0.67	0.1	0.5	27.3	90.3	9.6	64.6	0.43	-0.1	0.4	-19.6	76.5	-33.1	77.0
Crustal $PM_{2.5}$	110	4.96	2.33	-2.6	3.0	-53.1	60.5	-56.6	75.4	2.60	-2.3	2.8	-47.1	57.7	-49.2	71.1
Total $PM_{2.5}$	1909	18.49	17.06	-1.4	5.0	-7.7	27.3	-11.5	33.4	15.56	-2.7	6.0	-14.7	32.9	-19.3	40.4

Table 6A-10. Air quality model performance in the Southeastern (S) US (January 2002)

			January 2002 (Base emissions)							January 2002 (Adjusted emissions)						
Species	No. OBS	OBS conc.	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE
O_3	13	44.65	24.07	-20.6	21.3	-46.1	47.7	-65.0	66.7	32.30	-12.4	13.2	-27.8	29.5	-36.0	37.8
CO	1202	754.76	435.28	-319.5	414.7	-42.3	54.9	-44.0	65.4	557.39	-192.7	403.8	-25.7	53.8	-24.7	60.5
NO_2	1408	14.81	13.25	-1.6	6.0	-10.5	40.8	-15.0	49.9	8.71	-6.0	7.6	-40.8	51.5	-59.1	72.5
NO_x	818	28.87	19.39	-9.5	16.0	-32.8	55.3	-27.4	63.7	11.03	-17.6	19.0	-61.4	66.3	-74.0	86.4
NO_y	49	11.24	10.41	-0.8	5.7	-7.4	51.1	26.0	82.3	5.74	-4.9	7.4	-46.2	69.3	-7.0	111.6
SO_2	1367	6.23	6.25	0.0	3.9	0.3	62.8	-2.4	63.4	4.42	-1.7	3.6	-27.9	58.6	-31.5	67.3
HCHO	20	2.19	1.27	-0.9	1.0	-42.3	45.1	-46.1	51.8	2.06	-0.1	0.8	-6.3	38.3	-4.1	36.3
CCHO	21	0.98	0.55	-0.4	0.5	-44.2	48.4	-42.3	52.8	0.94	0.0	0.4	-4.6	42.8	5.7	40.9
Acetone	21	1.13	0.44	-0.7	0.7	-60.9	66.4	-68.3	82.5	0.63	-0.5	0.6	-44.2	54.4	-37.6	61.9
Ethylene	191	6.33	1.23	-5.1	5.2	-80.6	81.8	-92.8	100.0	2.66	-3.6	4.8	-57.6	77.4	-42.6	82.2
Isoprene	144	0.04	0.02	0.0	0.0	-57.8	98.6	23.5	101.4	0.04	0.0	0.1	-8.4	125.8	68.6	122.7
ALK1	206	17.15	1.09	-16.1	16.1	-93.7	93.7	-168.4	168.4	1.99	-15.2	15.2	-88.4	88.4	-147.3	147.5
ALK2	121	16.26	1.81	-14.4	14.4	-88.9	88.9	-148.9	148.9	3.16	-13.2	13.2	-80.7	80.7	-121.0	121.0
ALK3	206	12.16	2.32	-9.8	9.9	-81.0	81.4	-122.0	123.3	4.83	-7.2	8.7	-59.9	72.5	-77.6	92.7
ALK4	164	10.25	2.20	-8.0	8.2	-78.5	79.8	-105.7	111.5	4.49	-5.8	7.4	-56.1	72.7	-60.6	86.5
ALK5	154	2.26	1.73	-0.5	2.0	-23.4	89.3	15.0	85.1	3.40	1.1	3.2	46.4	136.3	56.1	101.6
ARO1	197	2.69	1.02	-1.7	1.8	-62.2	67.9	-72.4	81.5	2.14	-0.6	2.1	-20.9	77.7	-18.8	68.9
ARO2	75	2.81	0.92	-1.9	2.1	-67.1	73.7	-52.0	70.8	1.76	-1.0	2.3	-37.3	83.5	-8.5	69.8
OLE1	201	6.67	0.44	-6.2	6.2	-93.4	93.4	-149.4	149.5	0.94	-5.7	5.8	-85.8	87.8	-114.5	120.4
OLE2	197	0.79	0.39	-0.4	0.7	-51.3	94.9	0.3	116.7	0.81	0.0	1.0	1.5	121.1	35.0	112.6
TNMOC	136	53.46	18.25	-35.2	35.8	-65.9	66.9	-84.1	86.7	38.66	-13.7	29.0	-26.1	55.3	-26.2	60.2
SO_4^{2-} $PM_{2.5}$	378	2.42	2.46	0.0	1.2	1.9	51.0	4.9	49.4	2.13	-0.2	1.1	-10.3	45.0	-5.9	45.7
NO_3^- $PM_{2.5}$	331	1.20	5.04	3.8	4.0	319.6	331.3	84.9	123.8	2.09	0.9	1.3	74.4	110.3	9.2	96.8
NH_4^+ $PM_{2.5}$	271	1.07	2.50	1.4	1.5	132.7	144.0	67.0	82.7	1.38	0.3	0.7	30.9	63.9	20.0	59.8
OC $PM_{2.5}$	376	3.25	1.39	-1.9	2.0	-57.2	60.5	-68.4	77.9	3.07	-0.2	1.6	-5.9	49.7	-3.4	53.0
EC $PM_{2.5}$	376	0.65	0.60	-0.1	0.4	-8.0	64.3	-12.3	58.4	0.46	-0.2	0.4	-30.6	63.0	-37.8	66.0
Crustal $PM_{2.5}$	183	2.34	3.19	0.8	2.1	36.0	90.9	27.7	78.5	2.74	0.4	1.9	19.0	81.4	17.2	75.1
Total $PM_{2.5}$	2387	10.69	15.91	5.2	7.3	48.8	68.1	28.8	52.6	14.00	3.4	5.8	32.1	55.0	19.3	46.6

Table 6A-11. Air quality model performance in Georgia (G) (July 2001)

			July 2001 (Base emissions)							July 2001 (Adjusted emissions)						
Species	No. OBS	OBS conc.	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE
O_3	91	47.55	68.71	21.2	21.6	44.5	45.4	36.3	37.0	62.25	14.2	19.3	29.5	40.2	22.6	34.2
CO	37	460.21	472.31	12.1	216.4	2.6	47.0	8.3	53.3	397.62	-40.8	186.3	-9.3	42.5	-1.9	46.9
NO_2	50	9.57	12.54	3.0	5.8	31.0	60.9	24.0	73.1	7.81	-1.4	5.5	-14.8	59.9	-8.3	72.8
NO_x	50	15.64	13.35	-2.3	10.9	-14.7	70.0	11.0	86.0	8.34	-6.3	11.1	-43.2	75.3	-17.6	86.6
NO_y	37	25.00	24.92	-0.1	15.3	-0.3	61.1	21.7	64.0	15.61	-7.9	14.5	-33.7	61.8	-13.2	61.4
SO_2	100	1.90	3.40	1.5	2.1	78.5	111.8	53.4	76.6	2.09	0.1	1.3	5.3	66.4	8.6	58.7
HCHO	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CCHO	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Acetone	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Ethylene	15	1.20	1.42	0.2	0.8	18.6	69.2	9.0	69.2	0.64	-0.5	0.6	-44.9	50.9	-56.9	65.4
Isoprene	25	1.25	2.85	1.6	1.6	128.6	131.8	73.2	75.7	0.78	-0.4	0.5	-34.6	40.6	-43.5	51.6
ALK1	26	1.91	0.93	-1.0	1.0	-51.2	52.3	-71.2	72.8	0.54	-1.5	1.5	-73.0	73.0	-111.6	111.6
ALK2	26	1.62	4.03	2.4	2.4	149.1	149.1	93.5	93.5	1.45	-0.2	0.7	-9.7	40.5	3.5	43.7
ALK3	26	0.91	3.16	2.3	2.4	248.2	261.8	93.2	101.5	1.47	0.6	1.0	63.9	105.7	35.2	71.6
ALK4	26	2.35	2.46	0.1	1.6	4.9	68.7	-2.3	65.8	1.12	-1.2	1.4	-51.5	60.3	-66.7	79.4
ALK5	23	0.83	1.32	0.5	0.9	59.7	114.3	44.9	82.7	0.61	-0.2	0.6	-25.5	69.6	-18.0	77.4
ARO1	25	1.27	1.22	0.0	0.7	-3.6	57.3	-0.5	61.4	0.56	-0.7	0.8	-55.7	61.2	-67.6	86.0
ARO2	8	2.61	0.68	-1.9	1.9	-74.1	74.1	-128.8	128.8	0.32	-2.3	2.3	-87.7	87.7	-161.3	161.3
OLE1	22	0.49	0.56	0.1	0.4	14.7	85.5	-9.0	78.7	0.25	-0.2	0.3	-48.0	59.3	-73.4	86.4
OLE2	6	0.17	0.19	0.0	0.1	13.0	82.7	-4.1	63.8	0.07	-0.1	0.1	-58.0	74.7	-85.4	101.4
TNMOC	26	24.16	54.89	30.7	38.3	127.1	158.4	91.3	102.9	22.62	0.0	15.0	0.1	66.4	24.8	64.8
SO_4^{2-}-$PM_{2.5}$	48	8.73	12.85	4.1	4.2	47.3	48.0	41.5	42.3	8.58	-0.2	2.7	-2.4	30.5	-3.4	38.1
NO_3^- $PM_{2.5}$	48	0.52	0.83	0.3	0.7	58.2	126.7	1.5	93.2	0.49	0.0	0.4	-4.3	87.6	-42.6	97.1
NH_4^+ $PM_{2.5}$	46	3.30	4.55	1.2	1.9	37.7	56.5	44.0	56.8	3.09	-0.3	1.4	-8.0	41.3	4.9	51.2
OC $PM_{2.5}$	54	7.79	3.19	-4.6	4.9	-59.1	63.1	-71.1	79.7	5.73	-2.3	4.1	-28.8	50.9	-26.2	55.9
EC $PM_{2.5}$	53	0.99	1.24	0.3	0.5	25.3	52.7	20.3	47.2	0.88	-0.1	0.4	-12.0	43.0	-15.0	50.0
Crustal $PM_{2.5}$	23	8.02	5.38	-2.6	3.6	-32.9	44.8	-22.9	54.3	6.06	-1.7	2.9	-21.5	37.6	-11.3	44.7
Total $PM_{2.5}$	168	24.86	25.32	0.5	6.8	1.8	27.5	1.2	30.2	24.36	-0.8	7.7	-3.0	30.5	-4.0	34.2

Table 6A-12. Air quality model performance in Georgia (G) (January 2002)

			January 2002 (Base emissions)							January 2002 (Adjusted emissions)						
Species	No. OBS	OBS conc.	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE	MODEL conc	MB [μg /ppb]	ME	NMB	NME [%]	MFB	MFE
O_3	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CO	25	833.38	628.18	-205.2	313.2	-24.6	37.6	-29.3	44.8	677.84	-149.8	313.9	-18.1	37.9	-23.4	44.0
NO_2	65	17.19	16.65	-0.5	7.0	-3.1	40.4	-2.3	50.5	8.54	-9.1	9.7	-51.6	54.7	-71.3	81.3
NO_x	65	42.22	24.72	-17.5	26.8	-41.4	63.5	-21.5	74.2	9.78	-34.0	34.8	-77.6	79.7	-93.6	109.2
NO_y	24	50.29	31.48	-18.8	32.2	-37.4	64.1	-39.0	76.4	13.72	-39.1	39.8	-74.0	75.4	-103.0	106.3
SO_2	91	2.73	3.97	1.2	2.2	45.7	79.8	42.4	65.6	2.52	-0.6	1.8	-20.2	57.4	-8.2	57.3
HCHO	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CCHO	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Acetone	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Ethylene	10	1.91	1.52	-0.4	0.9	-20.4	45.6	-28.0	48.4	0.87	-1.0	1.0	-54.4	54.4	-71.5	71.5
Isoprene	1	0.08	0.00	-0.1	0.1	-95.1	95.1	-181.4	181.4	0.01	-0.1	0.1	-91.1	91.1	-167.4	167.4
ALK1	10	3.67	1.30	-2.4	2.4	-64.5	64.5	-96.7	96.7	0.93	-2.7	2.7	-74.7	74.7	-119.6	119.6
ALK2	7	4.20	1.93	-2.3	2.3	-54.0	54.0	-73.2	73.2	2.21	-2.0	2.0	-47.4	47.4	-62.7	62.7
ALK3	10	2.83	3.67	0.8	1.2	29.6	42.0	23.4	32.6	2.11	-0.7	0.9	-25.5	31.9	-28.3	36.6
ALK4	10	2.86	3.21	0.4	1.2	12.4	40.5	9.5	33.6	1.72	-1.1	1.1	-39.8	39.8	-47.2	47.2
ALK5	10	0.66	1.78	1.1	1.1	170.9	170.9	93.3	93.3	1.00	0.3	0.4	52.4	64.1	45.3	54.4
ARO1	10	0.95	1.75	0.8	0.8	85.4	85.4	65.1	65.1	1.00	0.1	0.4	5.7	43.2	14.5	51.8
ARO2	8	0.88	1.37	0.5	0.7	54.5	83.9	35.5	65.0	0.62	-0.3	0.4	-29.7	46.8	-28.6	59.2
OLE1	10	0.58	0.79	0.2	0.3	35.4	48.7	25.4	33.9	0.30	-0.3	0.3	-47.9	47.9	-60.1	60.1
OLE2	10	0.29	0.50	0.2	0.3	75.1	105.8	49.2	82.7	0.18	-0.1	0.2	-37.6	55.5	-33.9	72.2
TNMOC	8	28.09	21.97	-6.1	11.0	-21.8	39.1	-31.7	46.4	16.45	-11.6	12.3	-41.4	43.6	-52.0	55.2
SO_4^{2-} $PM_{2.5}$	64	2.26	2.87	0.6	1.0	27.4	43.3	22.4	38.4	2.33	0.2	0.8	7.8	35.4	7.1	34.4
NO_3^- $PM_{2.5}$	64	1.72	6.79	5.1	5.3	295.1	307.2	90.8	116.8	2.49	0.7	1.4	39.0	78.0	17.1	74.6
NH_4^+ $PM_{2.5}$	55	0.94	3.21	2.3	2.4	241.3	257.3	102.6	114.2	1.65	0.8	1.0	84.9	113.1	66.1	87.8
OC $PM_{2.5}$	63	4.05	2.37	-1.7	1.9	-41.4	46.1	-43.0	51.8	4.33	0.1	1.7	3.0	40.7	9.8	41.7
EC $PM_{2.5}$	59	0.73	1.13	0.4	0.7	55.2	94.1	43.5	75.6	0.67	-0.1	0.5	-10.7	67.4	2.3	73.3
Crustal $PM_{2.5}$	39	3.19	5.89	2.7	3.4	84.8	106.6	77.9	88.6	5.60	2.5	3.2	82.4	104.7	77.3	88.2
Total $PM_{2.5}$	206	12.28	23.16	10.9	12.5	88.6	101.4	55.9	68.3	19.26	6.9	8.7	56.4	70.4	41.6	54.6

Acknowledgement

This research was supported by the U.S. Environmental Protection Agency under Agreements RD82897602, RD83107601, and RD83096001.

References

Aldrin, M. (1997). Length modified ridge regression. *Computational Statistics & Data Analysis* **25**, 377-398.

Byun, D. W., and Ching, J. K. (1999). Science algorithms of the EPA Models-3 Community Multiscale Air Qualtiy (CMAQ) modeling system. *Atmospheric modeling division, National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711*.

Census (2005). Census Regions and Divisions of the United States. *http://www.census.gov/geo/www/us_regdiv.pdf*.

Chang, M. E., Hartley, D. E., Cardelino, C., and Chang, W. L. (1996). Inverse modeling of biogenic isoprene emissions. *Geophysical Research Letters* **23**, 3007-3010.

Frank, I. E., and Friedman, J. H. (1993). A Statistical View Of Some Chemometrics Regression Tools. *Technometrics* **35**, 109-135.

Gilliland, A. B., Dennis, R. L., Roselle, S. J., and Pierce, T. E. (2003). Seasonal NH3 emission estimates for the eastern United States based on ammonium wet concentrations and an inverse modeling method. *Journal Of Geophysical Research-Atmospheres* **108**.

Harley, R. A., Russell, A. G., McRae, G. J., Cass, G. R., and Seinfeld, J. H. (1993). Photochemical Modeling Of The Southern California Air-Quality Study. *Environmental Science & Technology* **27**, 378-388.

Hoerl, A. E., and Kennard, R. W. (1970). Ridge Regression - Biased Estimation For Nonorthogonal Problems. *Technometrics* **12**, 55-&.

Marquardt, D. W. (1963). An Algorithm For Least-Squares Estimation Of Nonlinear Parameters. *Journal Of The Society For Industrial And Applied Mathematics* **11**, 431-441.

Mendoza-Dominguez, A., and Russell, A. G. (2000). Iterative inverse modeling and direct sensitivity analysis of a photochemical air duality model. *Environmental Science & Technology* **34**, 4974-4981.

Mendoza-Dominguez, A., and Russell, A. G. (2001a). Emission strength validation using four-dimensional data assimilation: Application to primary aerosol and precursors to ozone and secondary aerosol. *Journal Of The Air & Waste Management Association* **51**, 1538-1550.

Mendoza-Dominguez, A., and Russell, A. G. (2001b). Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. *Atmospheric Environment* **35**, 2879-2894.

Menke, W. (1989). Geophysical Data Analysis: Discrete Inverse Theory, Revised Edition. *Academic Press*.

Muller, J. F., and Stavrakou, T. (2005). Inversion of CO and NOx emissions using the adjoint of the IMAGES model. *Atmospheric Chemistry And Physics* **5**, 1157-1186.

Odman, M. T., and Russell, A. G. (1991). Multiscale Modeling Of Pollutant Transport And Chemistry. *Journal Of Geophysical Research-Atmospheres* **96**, 7363-7370.

Pace, T. G. (2003). A conceptual model to adjust fugitive dust emissions to account for near source article removal in grid model applications. *US EPA. http://www.epa.gov/ttn/chief/emch/invent/statusfugdustemissions_082203.pdf.*

Petron, G., Granier, C., Khattatov, B., Yudin, V., Lamarque, J. F., Emmons, L., Gille, J., and Edwards, D. P. (2004). Monthly CO surface sources inventory based on the 2000-2001 MOPITT satellite data. *Geophysical Research Letters* **31**.

Placet, M., Mann, C. O., Gilbert, R. O., and Niefer, M. J. (2000). Emissions of ozone precursors from stationary sources: a critical review. *Atmos. Environ.* **34**, 2183-2204.

PSU/NCAR (2003). PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3. *Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research.*

US-EPA (2003). Clean air markets -- Data and maps. *http://cfpub.epa.gov/gdm/.*

US-EPA (2004). SMOKE User's manual. *http://cf.unc.edu/cep/empd/products/smoke*.

CHAPTER 7

CONCLUSIONS AND FUTURE RESEARCH

Conclusions

Elevated ambient particle levels affect human health, reduce visibility, deposit acid, and change Earth radiation balance. Because of the adverse health effect of particle concentrations, in 1997, the National Ambient Air Quality Standard (NAAQS) for $PM_{2.5}$ was promulgated. Metropolitan Atlanta, Georgia is one such area where $PM_{2.5}$ levels exceed the NAAQS standard. Strategies to control these fine particle levels should be proposed. This study estimated the controls required to meet the standard.

The proposed control strategies can be tested via air quality models, which have been used for air quality management and investigation for decades. For the models to be used in such applications, a thorough evaluation of these models is needed. This thesis carefully analyzed the major sources of errors in current air quality models and proposes ways to improve model performance. The findings of this research and analysis are summarized below.

Statistical analysis of $PM_{2.5}$ control strategy

Statistical analysis of $PM_{2.5}$ concentrations in metro Atlanta relied on measurements from four stations, Fort McPherson (FTM), South DeKalb (SDK), Tucker (TUC), and Jefferson Street (JST) in Atlanta. Emission reductions required to meet the annual $PM_{2.5}$ standard were 22% (FTM), 18% (SDK), 25% (JST), and 26% (TUC). $PM_{2.5}$ data in Atlanta were fit to lognormal distributions, and the distributional property

inherent in pollutant concentrations was used to estimate the variations of annual pollutant levels. The required emission reductions to meet the standard, when the temporal variations are considered, were 35% (TUC), 30% (FTM), 22% (SDK), and 40% (JST).

Uncertainty in air quality model evaluation for particulate matter due to spatial variation in pollutant concentrations

Air quality model performance is determined by the relative agreement between observed and simulated concentrations, though models predict volume-averaged concentrations, whereas monitors measure concentrations at a single point in space. The different spatial scales complicate the evaluation of the model performance and create uncertainty about the evaluation results if pollutant concentrations are spatially inhomogeneous. Spatial variability was assessed by comparing single point observed concentrations and interpolated concentrations derived from observations. The difference between interpolated and observed concentrations was comparable to the model error, suggesting that spatial variability should not be ignored in evaluating model performance. To decrease the influence of spatial variability on the model performance, the use of observations only from spatially representative stations has been suggested. The overall model performance of $PM_{2.5}$ using spatially representative stations apparently improved when compared with the performance of $PM_{2.5}$ from all stations.

Evaluation of fine particle number concentrations in CMAQ

Fine particle number concentrations predicted by CMAQ were evaluated using size-resolved particle number concentrations measured from Jan. 1, 1999, to Aug. 31, 2000 in Atlanta, Georgia. Simulated number concentrations were significantly underestimated, by a factor of as much as 1,000, whereas simulated mass concentrations were reasonably well reproduced. Causes of the incorrect prediction of particle number concentrations include that the assumed particle density is lower than the actual, the assumed size of primary particle emissions is too large, and the treatment of aerosols as three lognormal modes cannot accurately simulate nucleated particles that account for a large percentage of particle number concentrations. A more accurate prediction of particle number concentrations is expected when the above assumptions in CMAQ are modified.

Comparison between chemical mass balance receptor and CMAQ model $PM_{2.5}$ source apportionment

Source apportionment of particulate matter has been commonly performed using receptor models, but studies suggest that basic assumptions limit the accuracy and the extent of results. An alternative approach is the use of three-dimensional air quality models. A comparison was made between $PM_{2.5}$ sources apportioned from a receptor model and those from a source-based air quality model. The two results had a moderate discrepancy. Mass contributions calculated from the air quality model do not have enough monthly/day-of-week/hourly variations, whereas, those from the receptor models fluctuate more than the real variations because of measurement errors. Other sources of

discrepancy include different spatial scales, fundamental limitations, and uncertainties about the two models. Each model's strength can be used to help circumvent the other model's limitations.

Regional adjustment of emission strength via four-dimensional data assimilation

FDDA was used to adjust the emission strengths for ten sources: CO (total), NH3 (total), SO_2 (total), NO_x (area/mobile/nonroad), NO_x (point), VOC (area/mobile/nonroad/point), VOC (biogenic), POA (total), PEC (total), and PMFINE (total). Results showed that the adjustment was relatively mild for CO and SO_2 emissions. The adjusted NH_3, NO_x, PEC and PMFINE emissions were lower than the base level emissions. The emission-scaling factors for POA and VOC emissions were significantly different among regions. The different model performance among VOC species complicated the adjustment of VOC emissions. The difference in the emission adjustment between weekdays and weekends was minor in most cases.

Future Research

Representative monitoring sites for the regional air quality model

The representativeness of monitoring stations is important in epidemiological study and air quality model evaluation. The usage of observations and simulated concentrations, which has different spatial scales, generates uncertainly in model performance. Thus, an appropriate handling of this problem is necessary to improve the understanding of the model performance (see Chapter 3). The main goal of this study is

using spatial interpolation and multiple hypotheses testing techniques to evaluate current monitoring stations in the US in terms of their representativeness. Because three-dimensional Eulerian photochemical models generate volume-averaged concentrations, it is desirable to use volume-averaged measurements when evaluating models.

Three different interpolation techniques such as inverse distance square weighted interpolation, kriging, and spline, will be used to obtain the predicted volume-averaged measurements. The difference between observed and predicted values for each station implies how well this station represents the surrounding area. In other words, if the difference between observed and predicted value is small in statistical sense, the representativeness of the corresponding stations is supported. In order to provide a systematic threshold of the difference, we propose a multiple testing procedure based on the concept of the False Discovery Rate (FDR) (Benjamini and Hochberg, 1995). The FDR is defined to be the expected proportion of false positives among all the hypotheses rejected

$$FDR \equiv E\left[\frac{V}{R}\right], \qquad (7\text{-}1)$$

where R is the number of rejected null hypotheses, and V is the number of falsely rejected null hypotheses. Benjamini and Hochberg proved that an ordered p-value procedure controls the specified FDR (reference). The details of this procedure will be presented in our paper. Our null hypothesis states that the difference between observed and predicted values is small (i.e., the monitoring stations is well representative). Assuming that the number of monitoring stations is 1,206 (the number of $PM_{2.5}$ monitors, see Chapter 3), the number of hypothesis is 1,206. Applying the FDR procedure to all 1,206 hypotheses simultaneously gives a list of monitoring stations well representative. The results of this

study can provide useful information to determine appropriate inputs of air quality model evaluation by selecting qualified monitoring stations. Moreover, this study can provide a guideline to future areas to be monitored as well.

Multivariate statistical analyses of the fine particle number concentrations

Particles in the atmosphere change climate through cloud-aerosol interactions and the elevated levels of the ambient particle number adversely affect human health (Donaldson et al., 1998; Penttinen et al., 2001; Woo et al., 2001a; Woo et al., 2001b). Currently, the new particle formation by nucleation are not well understood, accordingly new particle number concentrations are not accurately simulated in the air quality model (see Chapter 6). The goal of this study is to fully investigate the factors affecting the particle nucleation process. Previous studies revealed that low temperature and higher nucleation precursor concentrations are key factors for the nucleation process (Woo et al., 2001a). However, those studies did not predict the occurrence of the nucleated particle number concentrations with those factors. In this study, multivariate statistical analyses will be employed to produce the rules that predict the number of particle as well as classify different types of day (i.e., high nucleation days and low nucleation days) based on various factors such as meteorology and gas phase precursor concentrations. Further, variable selection methods can identify the important factors that affect the particle nucleation process. The result of this study will provide a better understanding of the particle formation via nucleation, and can be also incorporated into the air quality models to improve the accuracy of the particle simulation.

Evaluation of the Potential Source Contribution Function (PSCF) analysis

The Potential Source Contribution Function (PSCF) is the receptor model that combines the particle data with air parcel back trajectories to identify potential source areas and the preferred pathways that give rise to the observed high particle concentrations(Ashbaugh et al., 1985; Zeng and Hopke, 1989). The PSCF value is defined as the conditional probability that a specified cell is associated with the "polluted" events given that trajectory endpoints fall into a specified cell. This evaluation of this method is usually done by comparing the PSCF values with the ground level emission map (Antoine, 2002; Begum et al., 2005). The emission map shows only the spatial distribution of emissions, and areas of the strong emissions are not necessarily the source locations for the specified pollutant in the receptor. This way of evaluation does not always carry the information whether the resulting probability field provides accurate estimates of the source locations or not. The goal of this study is to evaluate the PSCF method more comprehensively. The source locations and strengths investigated via the regional sensitivity analysis of the three-dimensional air quality model will be used as the reference for the evaluation. To make the result of PSCF be consistent with that of air quality modeling, the PSCF values will be calculated using the meteorology and pollutant levels used in, or derived by the air quality model.

References

Abdel-Aziz, A., and Frey, H. C. (2004). Propagation of uncertainty in hourly utility NOx emissions through a photochemical grid air quality model: A case study for the Charlotte, NC, modeling domain. *Environmental Science & Technology* **38**, 2153-2160.

Adrian, R. J., and Orloff, K. L. (1977). Laser Anemometer Signals - Visibility Characteristics And Application To Particle Sizing. *Applied Optics* **16**, 677-684.

Aldrin, M. (1997). Length modified ridge regression. *Computational Statistics & Data Analysis* **25**, 377-398.

Ames, R. B., and Malm, W. C. (2001). Chemical species' contributions to the upper extremes of aerosol fine mass. *Atmospheric Environment* **35**, 5193-5204.

Antoine, M. (2002). Transport characteristics and regional source assessment of PM2.5 in Atlanta: Cluster analysis and potential source contribution function analysis. *M.S. Thesis, Georgia Institute of Technology, Atlanta Georgia*.

Arnold, J. R., Dennis, R. L., and Tonnesen, G. S. (2003). Diagnostic evaluation of numerical air quality models with specialized ambient observations: testing the Community Multiscale Air Quality modeling system (CMAQ) at selected SOS 95 ground sites. *Atmospheric Environment* **37**, 1185-1198.

Ashbaugh, L. L., Malm, W. C., and Sadeh, W. Z. (1985). A Residence Time Probability Analysis Of Sulfur Concentrations At Grand-Canyon-National-Park. *Atmospheric Environment* **19**, 1263-1270.

Baumann, K., Chang, M. E., Russell, A. G., and Edgerton, E. S. (2005). Seasonal Variability of Organic Mass Contribution to PM2.5 within Metro Atlanta are Further Downwind. *AAAR international specialty conference -- Particulate Matter Supersites Program and Related Studies, Atlanta, Georgia*.

Begum, B. A., Kim, E., Jeong, C. H., Lee, D. W., and Hopke, P. K. (2005). Evaluation of the potential source contribution function using the 2002 Quebec forest fire episode. *Atmospheric Environment* **39**, 3719-3724.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of Royal Statistical Society* **B 57**, 289-300.

Bhave, P., and Binkowski, F. S. (2005). Private communication.

Bhave, P. V., Fergenson, D. P., Prather, K. A., and Cass, G. R. (2001). Source apportionment of fine particulate matter by clustering single-particle data: Tests of receptor model accuracy. *Environmental Science & Technology* **35**, 2060-2072.

Blanchard, C. L., Carr, E. L., Collins, J. F., Smith, T. B., Lehrman, D. E., and Michaels, H. M. (1999). Spatial representativeness and scales of transport during the 1995 Integrated Monitoring Study in California's San Joaquin Valley. *Atmospheric Environment* **33**, 4775-4786.

Boylan, J. W. (2004). PM model performance metrics, goals, and criteria. *Proceeding of 2004 Models-3 conference, Research Triangle, NC.*

Boylan, J. W., Odman, M. T., Wilkinson, J. G., Russell, A. G., Doty, K. G., Norris, W. B., and McNider, R. T. (2002). Development of a comprehensive, multiscale "one-atmosphere" modeling system: application to the Southern Appalachian Mountains. *Atmospheric Environment* **36**, 3721-3734.

Boylan, J. W., Odman, M. T., Wilkinson, J. G., Russell, A. G., Doty, K. G., Norris, W. B., and McNider, R. T. (2005). Integrated assessment modeling of atmospheric pollutants in the Southern Appalachian Mountains. Part I: Hourly and seasonal ozone. *Journal Of The Air & Waste Management Association* **55**, 1019-1030.

Brauer, M., Ebelt, S. T., Fisher, T. V., Brumm, J., Petkau, A. J., and Vedal, S. (2001). Exposure of chronic obstructive pulmonary disease patients to particles: Respiratory and cardiovascular health effects. *Journal Of Exposure Analysis And Environmental Epidemiology* **11**, 490-500.

Butler, A. J. (2000). Temporal and spatial analysis of PM2.5 mass and composition in Atlanta. *Ph. D. Thesis, School of Civil and Environmental Engineering, Georgia Institute of Technology*.

Butler, A. J., Andrew, M. S., and Russell, A. G. (2003). Daily sampling of PM2.5 in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. *Journal Of Geophysical Research-Atmospheres* **108**.

Byun, D. W., and Ching, J. K. (1999). Science algorithms of the EPA Models-3 Community Multiscale Air Qualtiy (CMAQ) modeling system. *Atmospheric modeling division, National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711*.

Carter, W. (2000). Documentation of the SAPRC-99 chemical mechanism for VOC reactivity assessment. *Final report to California air resources board contract 92-329 and contract 95-308. Air pollution research center and college of engineering, Center for Environmental Research and Technology, University of California, Riverside, California.*

Carter, W. P. L. (1996). Condensed atmospheric photooxidation mechanisms for isoprene. *Atmospheric Environment* **24A**, 481-515.

Carter, W. P. L., Luo, D., and Malkina, I. L. (1997). Environmental Chamber Studies for Devalopment of an Updated Photochemical Mechanism for Reactivity Assessment. *Final Report for California Air Resources Board Contract No. 92-345, Coordniating Research Council, Inc., Project M-9 and National Renewable Energy Laboratory, Contract ZF-2-12252-07*.

Census (2005). Census Regions and Divisions of the United States. *http://www.census.gov/geo/www/us_regdiv.pdf*.

Chan, C. C., and Hwang, J. S. (1996). Site representativeness of urban air monitoring stations. *Journal Of The Air & Waste Management Association* **46**, 755-760.

Chang, J. S., Brost, R. A., Isaksen, I. S. A., Madronich, S., Middleton, P., Stockwell, W. R., and Waleck, C. J. (1987). A three-dimensional Eulerian acid deposition model: physical concepts and formulation. *Journal Of Geophysical Research-Atmospheres* **92**, 14,681-14,700.

Chang, M. E., Hartley, D. E., Cardelino, C., and Chang, W. L. (1996). Inverse modeling of biogenic isoprene emissions. *Geophysical Research Letters* **23**, 3007-3010.

Chock, D. P., and Sluchak, P. S. (1986). Estimating Extreme Values Of Air-Quality Data Using Different Fitted Distributions. *Atmospheric Environment* **20**, 989-993.

Chow, J. C., Frazier, C., A., and Watson, J. G. (1988). A survey of existing fugitive/area source characterization methods for receptor modeling. In: Hoffnagle, G. F.; Lebowitz, M. (Eds.). *Proceedings, Particulate matter/Fugitive Dusts: Measurements and Control in Western Arid Regions. Air Pollution Control Association, Pittsburgh, PA.*, 176-189.

Chow, J. C., Watson, J. G., Kuhns, H., Etyemezian, V., Lowenthal, D. H., Crow, D., Kohl, S. D., Engelbrecht, J. P., and Green, M. C. (2004). Source profiles for industrial, mobile, and area sources in the Big Bend Regional Aerosol Visibility and Observational study. *Chemosphere* **54**, 185-208.

Christensen, W. F. (2004). Chemical mass balance analysis of air quality data when unknown pollution sources are present. *Atmospheric Environment* **38**, 4305-4317.

Cooper, J. A. (1981). **Determination of Source Contributions to Fine and Coarse Suspended Particulate Levels in Petersville, Alabama**. *Report to Tennessee Valley Authority by NEA, Inc.*

Core, J. E., Cooper, J. A., Hanrahan, P. L., and Cox, W. M. (1982). Particulate Dispersion Model Evaluation - A New Approach Using Receptor Models. *Journal Of The Air Pollution Control Association* **32**, 1142-1147.

de Nevers, N., Neligan, R. E., and Slater, H. H. (1977). Air Pollution. *Academic Press, New York* **vol. 5**.

Demster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society (B)* **39**, 1-38.

Dennis, R. L. (1998). The environmental protection agency's third generation air quality modeling system: an overall perspective. *Proceedings of the American Meteorological Society 78th Annual Meeting, Phoenix, AZ, Jan. 11-16, 1998. 255-258.*

Dockery, D. W., and Pope, C. A. (1994). Acute Respiratory Effects of Particulate Air-Pollution. *Annual Review of Public Health* **15**, 107-132.

Dockery, D. W., Pope, C. A., Xu, X. P., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., and Speizer, F. E. (1993). An Association Between Air-Pollution And Mortality In 6 United-States Cities. *New England Journal Of Medicine* **329**, 1753-1759.

Donaldson, K., Li, X. Y., and MacNee, W. (1998). Ultrafine (nanometre) particle mediated lung injury. *Journal Of Aerosol Science* **29**, 553-560.

Donaldson, K., Stone, V., Seaton, A., and MacNee, W. (2001). Ambient particle inhalation and the cardiovascular system: Potential mechanisms. *Environmental Health Perspectives* **109**, 523-527.

Dunker, A. M. (1981). Efficient Calculation Of Sensitivity Coefficients For Complex Atmospheric Models. *Atmospheric Environment* **15**, 1155-1161.

Eatough, D. J., Long, R. W., Modey, W. K., and Eatough, N. L. (2003). Semi-volatile secondary organic aerosol in urban atmospheres: meeting a measurement challenge. *Atmospheric Environment* **37**, 1277-1292.

Ebelt, S. T., Petkau, A. J., Vedal, S., Fisher, T. V., and Brauer, M. (2000). Exposure of chronic obstructive pulmonary disease patients to particulate matter: Relationships between personal and ambient air concentrations. *Journal Of The Air & Waste Management Association* **50**, 1081-1094.

Eder, B., Yu, S., and Dennis, R. (2003). An Evaluation of the Models-3 CMAQ Aerosol Module. *PM 2003 Meeting -- Particulate Matter: Atmospheric Sciences, Exposure and the Forth Colloquium on PM and Human Health, Pittsburgh, PA.*

Elleman, R. A., Kotchenruther, R. A., Covert, D. S., Mass, C. F., and J., C. (2004). CMAQ aerosol number and mass evaluation for Pacific Northwest. *Models-3 User's workshop, Research Triangle Park, North Carolina.*

Elleman, R. A., Kotchenruther, R. A., Covert, D. S., Mass, C. F., and J., C. (2005). CMAQ aerosol number and mass evaluation for Pacific Northwest. *7th conference on atmospheric chemistry, The 85th AMS annual meeting, San Diego, CA.*

Emery, C., Tai, E., and Yarwood, G. (2001). Enhanced meteorological modeling and performance evaluation for two texas episodes. *Report to the Texas Natural Resources Conservation Commission, prepared by ENVIRON, International Corp, Novato, CA.*

Faxvog, F. R., and Roessler, D. M. (1978). Carbon Aerosol Visibility Vs Particle-Size Distribution. *Applied Optics* **17**, 2612-2616.

Fletcher, R. (2003). Practical methods of optimization. *Chichester, Wiley.*

Frank, I. E., and Friedman, J. H. (1993). A Statistical View Of Some Chemometrics Regression Tools. *Technometrics* **35**, 109-135.

Friedlander, S. (1973). Chemical Element Balances And Identification Of Air-Pollution Sources. *Environmental Science & Technology* **7**, 235-240.

Gaydos, T. M., Stanier, C., and Pands, S. (2005). Modeling of in situ ultrafine atmospheric particle formation in the eastern United States. *Journal Of Geophysical Research-Atmospheres* **110**, doi:10.1029/2004JD004683.

Genikhovich, E. L., Ziv, A. D., Iakovleva, E. A., Palmgren, F., and Berkowicz, R. (2005). Joint analysis of air pollution in street canyons in St. Petersburg and Copenhagen. *Atmospheric Environment* **39**, 2747-2757.

Georgopoulos, P. G., and Seinfeld, J. H. (1982). Statistical Distributions Of Air Pollutant Concentrations. *Environmental Science & Technology* **16**, A401-&.

Gery, M. W., Whitten, G. Z., Killus, J. P., and Dodge, M. C. (1989). A photochemical kinetics mechanism for urban and regional scale computer modeling. *Journal Of Geophysical Research-Atmospheres* **94**, 12,925-12,956.

Gilliland, A. B., Dennis, R. L., Roselle, S. J., and Pierce, T. E. (2003). Seasonal NH3 emission estimates for the eastern United States based on ammonium wet concentrations and an inverse modeling method. *Journal Of Geophysical Research-Atmospheres* **108**.

Gold, D. R., Damokosh, A. I., Pope, C. A., Dockery, D. W., McDonnell, W. F., Serrano, P., Retama, A., and Castillejos, M. (1999). Particulate and ozone pollutant effects on the respiratory function of children in Southwest Mexico City (vol 10, pg 8, 1999). *Epidemiology* **10**, 470-470.

Goldberg, M. S., Burnett, R. T., Valois, M. F., Flegel, K., Bailar, J. C., Brook, J., Vincent, R., and Radon, K. (2003). Associations between ambient air pollution and daily mortality among persons with congestive heart failure. *Environmental Research* **91**, 8-20.

Gong, Q., and Demerjian, K. L. (1995). Hydrocarbon Losses On A Regenerated Nafion(R) Dryer. *Journal Of The Air & Waste Management Association* **45**, 490-493.

Gordon, G. E., Pierson, W. R., Daisey, J. M., Lioy, P. J., Cooper, J. A., Watson, J. G., and Cass, G. R. (1984). Considerations For Design Of Source Apportionment Studies. *Atmospheric Environment* **18**, 1567-1582.

Goswami, E., Larson, T., Lumley, T., and Liu, L. J. S. (2002). Spatial characteristics of fine particulate matter: Identifying representative monitoring locations in Seattle, Washington. *Journal Of The Air & Waste Management Association* **52**, 324-333.

Gray, H. A., and Cass, G. R. (1998). Source contributions to atmospheric fine carbon particle concentrations. *Atmospheric Environment* **32**, 3805-3825.

Hadley, A., and Toumi, R. (2003). Assessing changes to the probability distribution of sulphur dioxide in the UK using a lognormal model. *Atmospheric Environment* **37**, 1461-1474.

Hansen, D. A., Edgerton, E. S., Hartsell, B. E., Jansen, J. J., Kandasamy, N., Hidy, G. M., and Blanchard, C. L. (2003). The southeastern aerosol research and characterization study: Part 1-overview. *Journal Of The Air & Waste Management Association* **53**, 1460-1471.

Harley, R. A., Russell, A. G., McRae, G. J., Cass, G. R., and Seinfeld, J. H. (1993). Photochemical Modeling Of The Southern California Air-Quality Study. *Environmental Science & Technology* **27**, 378-388.

Hawthorne, S. B., Krieger, M. S., Miller, D. J., and Mathiason, M. B. (1989). Collection And Quantitation Of Methoxylated Phenol Tracers For Atmospheric-Pollution From Residential Wood Stoves. *Environmental Science & Technology* **23**, 470-475.

Held, T., Ying, Q., Kaduwela, A., and Kleeman, M. (2004). Modeling particulate matter in the San Joaquin Valley with a source-oriented externally mixed three-dimensional photochemical grid model. *Atmospheric Environment* **38**, 3689-3711.

Held, T., Ying, Q., Kleeman, M. J., Schauer, J. J., and Fraser, M. P. (2005). A comparison of the UCD/CIT air quality model and the CMB source-receptor model for primary airborne particulate matter. *Atmos. Environ. accepted for publication.*

Hildemann, L. M., Klinedinst, D. B., Klouda, G. A., Currie, L. A., and Cass, G. R. (1994). Sources Of Urban Contemporary Carbon Aerosol. *Environmental Science & Technology* **28**, 1565-1576.

Hildemann, L. M., Markowski, G. R., and Cass, G. R. (1991). Chemical-Composition Of Emissions From Urban Sources Of Fine Organic Aerosol. *Environmental Science & Technology* **25**, 744-759.

Hoerl, A. E., and Kennard, R. W. (1970). Ridge Regression - Biased Estimation For Nonorthogonal Problems. *Technometrics* **12**, 55-&.

Hogg, R. V., Craig, A. T., and McKean, J. W. (2005). Introduction to Mathematical Statistics. *Prentice Hall, New York.*

Hogrefe, C., Biswas, J., Lynn, B., Civerolo, K., Ku, J. Y., Rosenthal, J., Rosenzweig, C., Goldberg, R., and Kinney, P. L. (2004). Simulating regional-scale ozone climatology over the eastern United States: model evaluation results. *Atmospheric Environment* **38**, 2627-2638.

Hogrefe, C., Sistla, G., Zalewsky, E., Hao, W., and Ku, J. Y. (2003). An assessment of the emissions inventory processing systems EMS-2001 and SMOKE in grid-

based air quality models. *Journal Of The Air & Waste Management Association* **53**, 1121-1129.

Holland, D. M., and Fitzsimons, T. (1982). Fitting Statistical Distributions To Air-Quality Data By The Maximum-Likelihood Method. *Atmospheric Environment* **16**, 1071-1076.

Hu, Y., Cohan, D. S., Odman, M. T., and Russell, A. G. (2003). Air quality modeling of August 11-20, 2000 for the Fall Line Air Quality Study (FAQS). *Prepared for: Georgia Department of Natural Resources, Environmental Protection Division - Air Protection Branch, Atlanta, GA.*

Iwai, K., Mizuno, S., Miyasaka, Y., and Mori, T. (2005). Correlation between suspended particles in the environmental air and causes of disease among inhabitants: Cross-sectional studies using the vital statistics and air pollution data in Japan. *Environmental Research* **99**, 106-117.

Jang, C., Phillips, S., Possiel, N., Dolwick, P., B., T., Braverman, T., B., W., M., H., and Fox, T. (2004). Annual model simulations and evaluations of particulate matter and ozone over the continental United States using the Models-3/CMAQ System. *Proceedings of the Models-3 User's workshop, Research Triangle Park. North Carolina*.

Jarvis, C. H., and Stuart, N. (2001). A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part II: The interaction between number of guiding variables and the type of interpolation method. *Journal Of Applied Meteorology* **40**, 1075-1084.

Javitz, H. S., Watson, J. G., and Robinson, N. F. (1988). Performance of the chemical mass balance model with simulated local-scale aerosols. *Atmospheric Environment* **22**, 2309–2322.

Jiang, W. (2005). Differences Between CMAQ Fine Mode Particle and PM2.5 Concentrations and their Impact on Model Performance Evaluation in the Lower Fraser Valley. *Atmospheric Environment.*

Jiang, W., Giroux, E., Roth, H., and Yin, D. (2004). Evaluation of CMAQ PM results using size-resolved field measurement data: the particle diameter issue and its impact on model performance assessment. *ITM conference*.

Jiang, W., and Yin, D. (2001). Development and application of the PMx software package for converting CMAQ modal particulate matter results in size-resolved quantities. *Prepared for Program of Energy Research and Development, Natural Resources Canada. Report number: PET-1497-01S, 2001.*

Jun, M., and Stein, M. L. (2004). Statistical comparison of observed and CMAQ modeled daily sulfate levels. *Atmospheric Environment* **38**, 4427-4436.

Kao, A. S., and Friedlander, S. K. (1995). Frequency distribution of PM10 chemical components and their sources. *Environmental Science & Technology* **29**, 19-28.

Ke, L., Wang, F., Edgerton, E. S., Schauer, J. J., Russell, A. G., Cass, G. R., and Zheng, M. (2005). Source Apportionment of Organic Carbon in PM2.5 from the SEARCH study. *Atmospheric Environment, Submitted for publication*.

Kleeman, M. J., and Cass, G. R. (2001). A 3D Eulerian source-oriented model for an externally mixed aerosol. *Environmental Science & Technology* **35**, 4834-4848.

Koch, S. E., Desjardins, M., and Kocin, P. J. (1983). An Interactive Barnes Objective Map Analysis Scheme For Use With Satellite And Conventional Data. *Journal Of Climate And Applied Meteorology* **22**, 1487-1503.

Koehler, K. J., and Symanowski, J. T. (1995). Constructing multivariate distributions with specific marginal distributions. *Journal of multivariate analysis* **55**, 261-282.

Kruize, H., Hanninen, O., Breugelmans, O., Lebret, E., and Jantunen, M. (2003). Description and demonstration of the EXPOLIS simulation model: Two examples of modeling population exposure to particulate matter. *Journal Of Exposure Analysis And Environmental Epidemiology* **13**, 87-99.

Kulmala, M., Laaksonen, A., and Pirjola, L. (1998). Parameterizations for sulfuric acid/water nucleation rates. *Journal Of Geophysical Research-Atmospheres* **103**, 8301-8307.

Kulmala, M., Vehkamaki, H., Petajda, T., Dal Maso, M., Lauri, A., Kerminen, V. M., Birmili, W., and McMurry, P. H. (2004). Formation and growth rates of ultrafine atmospheric particles: a review of observations. *Journal Of Aerosol Science* **35**, 143-176.

Larsen, R. I. (1969). A new mathematical model of air pollutant concentration, averaging time and frequency. *Journal or the air pollution control association* **19**, 24.

Larson, T., Kalman, D., Wang, S., and Nothstein, G. (1989). Urban air toxics mitigation study, phase 1. *Technical report, Puget Sound Air Pollution Control Agency, Seattle, Washington, 1989.*

Latha, K. M., and Badarinath, K. V. S. (2005). Shortwave radiative forcing efficiency of urban aerosols - a case study using ground based measurements. *Chemosphere* **58**, 217-220.

Lin, J., Scheff, P. A., and Wadden, R. A. (1993). Development of a Two-Phase Receptor Model for NMHC and PM10 Air Pollution Sources in Chicago. *Proceeding of 86th annual meeting and exhibition of the Air & Waste Management Association, Denver, Colorado*.

Lu, H. C. (2002). The statistical characters of PM10 concentration in Taiwan area. *Atmospheric Environment* **36**, 491-502.

Lu, H. C., and Fang, G. C. (2002). Estimating the frequency distributions of PM10 and PM2.5 by the statistics of wind speed at Sha-Lu, Taiwan. *Science Of The Total Environment* **298**, 119-130.

Mannschreck, K., Klemp, D., Kley, D., Friedrich, R., Kuhlwein, J., Wickert, B., Matuska, P., Habram, M., and Slemr, F. (2002). Evaluation of an emission inventory by comparisons of modelled and measured emission ratios of individual HCs, CO and NOx. *Atmospheric Environment* **36**, S81-S94.

Manomaiphiboon, K., and Russell, A. G. (2003). Formulation of joint probability density function of velocity for turbulent flows: an alternative approach. *Atmospheric Environment* **37**, 4917-4925.

Marmur, A., Mulholland, J., and Russell, A. (2004). Air quality modeling of PM2.5 mass and composition in Atlanta: Results from a two-year simulation and implications for use in health studies. *Models-3 User's workshop, Research Triangle Park, North Carolina*.

Marmur, A., Park, S.-K., Mulholland, J., and Russell, A. G. (2005a). PM2.5 Source Apportionment Using Receptor and Source-oriented Models: Conceptual Differences and Implications for Health Studies. *Proceedings of an AAAR international specialty conference -- Particulate Matter Supersites Program and Related Studies, Atlanta, Georgia*.

Marmur, A., Unal, A., Russell, A. G., and Mulholland, J. A. (2005b). Optimization based source apportionment of $PM_{2.5}$: Incorporating Gas-to-Particle ratios. *Environ. Sci. Technol. accepted for publication*.

Marquardt, D. W. (1963). An Algorithm For Least-Squares Estimation Of Nonlinear Parameters. *Journal Of The Society For Industrial And Applied Mathematics* **11**, 431-441.

McKee, G. A., Wadden, R. A., and Scheff, P. A. (1990). Development of a Two-Phase Chemical Mass Balance Receptor Model. *Proceeding of the 83rd annual meeting and exhibition of the Air & Waste Management Association, Pittsburgh, Pennsylvania*.

McMurry, P. H. (1983). New Particle Formation In The Presence Of An Aerosol - Rates, Time Scales, And Sub-0.01 Mu-M Size Distributions. *Journal Of Colloid And Interface Science* **95**, 72-80.

McMurry, P. H., and Friedlander, S. K. (1979). New Particle Formation In The Presence Of An Aerosol. *Atmospheric Environment* **13**, 1635-1651.

McMurry, P. H., Wang, X., Park, K., and Ehara, K. (2002). The relationship between mass and mobility for atmospheric particles-A new technique for measuring particle density. *Aerosol Science and Technology* **36**, 227-238.

McMurry, P. H., and Woo, K. S. (2002). Size distributions of 3-100-nm urban Atlanta aerosols: Measurement and observations. *Journal Of Aerosol Medicine-Deposition Clearance And Effects In The Lung* **15**, 169-178.

McNair, L. A., Harley, R. A., and Russell, A. G. (1996). Spatial inhomogeneity in pollutant concentrations, and their implications for air quality model evaluation. *Atmospheric Environment* **30**, 4291-4301.

Mebust, M. R., Eder, B. K., Binkowski, F. S., and Roselle, S. J. (2003). Models-3 community multiscale air quality (CMAQ) model aerosol component - 2. Model evaluation. *Journal Of Geophysical Research-Atmospheres* **108**.

Mendoza-Dominguez, A., and Russel, A. G. (2001). Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. *Atmospheric Environment* **35**, 2879-2894.

Mendoza-Dominguez, A., and Russell, A. G. (2000). Iterative inverse modeling and direct sensitivity analysis of a photochemical air duality model. *Environmental Science & Technology* **34**, 4974-4981.

Mendoza-Dominguez, A., and Russell, A. G. (2001a). Emission strength validation using four-dimensional data assimilation: Application to primary aerosol and precursors to ozone and secondary aerosol. *Journal Of The Air & Waste Management Association* **51**, 1538-1550.

Mendoza-Dominguez, A., and Russell, A. G. (2001b). Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. *Atmospheric Environment* **35**, 2879-2894.

Menke, W. (1989). Geophysical Data Analysis: Discrete Inverse Theory, Revised Edition. *Academic Press.*

Metzger, K. B., Tolbert, P. E., Klein, M., Peel, J. L., Flanders, W. D., Todd, K., Mulholland, J. A., Ryan, P. B., and Frumkin, H. (2004). Ambient air pollution and cardiovascular emergency department visits. *Epidemiology* **15**, 46-56.

Meyer, M. B., Patashnick, H., Ambs, J. L., and Rupprecht, E. (2000). Development of a sample equilibration system for the TEOM continuous PM monitor. *Journal Of The Air & Waste Management Association* **50**, 1345-1349.

Morgan, M. G., and Henrion, M. (2003). A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. *Cambridge University Press.*

Mulhbaier, J. L., and Williams, R. L. (1982). Fireplaces, furnaces, and vehicles as emission sources of particulate carbons. *Particulate Carbon-Atmospheric Life Cycle, edited by G. T. Wolff*, 185-205.

Muller, J. F., and Stavrakou, T. (2005). Inversion of CO and NOx emissions using the adjoint of the IMAGES model. *Atmospheric Chemistry And Physics* **5**, 1157-1186.

Mysliwiec, M. J., and Kleeman, M. J. (2002). Source apportionment of secondary airborne particulate matter in a polluted atmospbere. *Environmental Science & Technology* **36**, 5376-5384.

Napelenok, S., Cohan, D. S., Russell, A. G., and Hu, Y. (2005). Implementation of the direct decoupled method (DDM) into the CMAQ model for aerosol species. *Proceedings of an AAAR international specialty conference -- Particulate Matter Supersites Program and Related Studies, Atlanta, Georgia*.

NCAR (2003). PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3. *Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research*.

Nenes, A. (1996). ISORROPIA v1.5: Reference manual. *http://nenes.eas.gatech.edu/ISORROPIA/Version1_5/ISORROPIA15Manual.pdf*.

Nenes, A., Pandis, S. N., and Pilinis, C. (1998). ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols. *Aquatic Geochemistry* **4**, 123-152.

Odman, M. T., and Russell, A. G. (1991). Multiscale Modeling Of Pollutant Transport And Chemistry. *Journal Of Geophysical Research-Atmospheres* **96**, 7363-7370.

Ott, W., Wallace, L., and Mage, D. (2000). Predicting particulate (PM10) personal exposure distributions using a random component superposition statistical model. *Journal Of The Air & Waste Management Association* **50**, 1390-1406.

Pace, T. G. (2003). A conceptual model to adjust fugitive dust emissions to account for near source article removal in grid model applications. *US EPA. http://www.epa.gov/ttn/chief/emch/invent/statusfugdustemissions_082203.pdf*.

Park, K., Kittelson, D. B., and McMurry, P. H. (2004a). Structural properties of diesel exhaust particles measured by transmission electron microscopy (TEM): Relationships to particle mass and mobility. *Aerosol Science And Technology* **38**, 881-889.

Park, K., Kittelson, D. B., Zachariah, M. R., and McMurry, P. H. (2004b). Measurement of inherent material density of nanoparticle agglomerates. *Journal Of Nanoparticle Research* **6**, 267-272.

Park, S.-K., Cobb, C. E., Wade, K., Mulholland, J., Hu, Y., and Russell, A. (2005a). Uncertainty in Air Qualiy Model Evaluation for Particulate Matter due to Spatial Variation in Pollutant Concentrations. *Atmospheric Environment, submitted for publication*.

Park, S.-K., Cobb, C. E., Wade, K., Mulholland, J., Hu, Y., and Russell, A. G. (2005b). Uncertainty in Air Quality Model Evaluation from Spatial Variation. *Submitted to Atmos. Environ.*

Park, S.-K., Cobb, C. E., Wade, K., Mulholland, J., Hu, Y., and Russell, A. G. (Submitted). Uncertainty in Air Quality Model Evaluation from Spatial Variation. *Atmospheric Environment*.

Park, S.-K., Ke, L., B., Y., G., R. A., and Zheng, M. (2005c). Source apportionment of PM2.5 using a three-dimensional air quality model and a receptor model. *Proceedings of an AAAR international specialty conference -- Particulate Matter Supersites Program and Related Studies, Atlanta, Georgia*.

Park, S.-K., Marmur, A., Ke, L., Yan, B., Russell, A. G., and Zheng, M. (submitted for publication). Source apportionment of PM2.5: Inter-comparison between receptor and emission-based models. *Environmental Science and Technology*.

Park, S.-K., Marmur, A., Ke, L., Yan, B., Zheng, M., and Russell, A. G. (2005d). Comparison between Chamical Mass Balance Receptor and CMAQ MODEL PM2.5 Source Apportionment. *Submitted to Environmental Science and Technology*.

Park, S.-K., and Russell, A. G. (2003a). Sensitivity of PM 2.5 to emissions in the Southeast. *Proceedings of the Models-3 User's workshop, Research Triangle Park, North Carolina.*

Park, S.-K., and Russell, A. G. (2003b). Sensitivity of PM 2.5 to emissions in the Southeast. *Proceedings of the Models-3 User's workshop, Research Triangle Park, North Carolina.*

Patashnick, H., and Rupprecht, E. G. (1991). Continuous Pm-10 Measurements Using The Tapered Element Oscillating Microbalance. *Journal Of The Air & Waste Management Association* **41**, 1079-1083.

Peel, J., Tolbert, P., Klein, M., Metzger, K., Flanders, W. D., Todd, K., Mulholland, J., Ryan, P. B., and Frumkin, H. (2002). Ambient air pollution and respiratory emergency department visits in Atlanta, August 1998-August 2000 (ARIES/SOPHIA). *Epidemiology* **13**, S124-S124.

Penttinen, P., Timonen, K. L., Tittanen, P., Mirme, A., Ruuskanen, J., and Pekkanen, J. (2001). Number concentration and size of particles in urban air: Effects on Spirometric lung function in adult asthmatic subjects. *Environmental Health Perspectives* **109**, 319-323.

Petron, G., Granier, C., Khattatov, B., Yudin, V., Lamarque, J. F., Emmons, L., Gille, J., and Edwards, D. P. (2004). Monthly CO surface sources inventory based on the 2000-2001 MOPITT satellite data. *Geophysical Research Letters* **31**.

Placet, M., Mann, C. O., Gilbert, R. O., and Niefer, M. J. (2000a). Emissions of ozone precursors from stationary sources: a critical review. *Atmospheric Environment* **34**, 2183-2204.

Placet, M., Mann, C. O., Gilbert, R. O., and Niefer, M. J. (2000b). Emissions of ozone precursors from stationary sources: a critical review. *Atmos. Environ.* **34**, 2183-2204.

Pope, C. A., Thun, M. J., Namboodiri, M. M., Dockery, D. W., Evans, J. S., Speizer, F. E., and Heath, C. W. (1995). Particulate Air-Pollution As A Predictor Of Mortality In A Prospective-Study Of Us Adults. *American Journal Of Respiratory And Critical Care Medicine* **151**, 669-674.

Pope, C. A., Verrier, R. L., Lovett, E. G., Larson, A. C., Raizenne, M. E., Kanner, R. E., Schwartz, J., Villegas, M., Gold, D. R., and Dockery, D. W. (1999). Heart rate variability associated with particulate air pollution. *American Heart Journal* **138**, 890-899.

PSU/NCAR (2003). PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3. *Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research.*

Redner, A. R., and Walker, H. F. (1984). Mixture densities, maximum likelihood and the EM algorithm. *SIAM review* **26**, 195-239.

Restrepo, C., Zimmerman, R., Thurston, G., Clemente, J., Gorczynski, J., Zhong, M. H., Blaustein, M., and Chen, L. C. (2004). A comparison of ground-level air quality data with New York State Department of Environmental Conservation monitoring stations data in South Bronx, New York. *Atmospheric Environment* **38**, 5295-5304.

Roberts, D. L., and Friedlander, S. K. (1980). Sulfur-Dioxide Transport Through Aqueous-Solutions.2. Experimental Results And Comparison With Theory. *Aiche Journal* **26**, 602-610.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., and Cass, G. R. (1994). Sources Of Fine Organic Aerosol.6. Cigarette-Smoke In The Urban Atmosphere. *Environmental Science & Technology* **28**, 1375-1388.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1993a). Sources Of Fine Organic Aerosol.2. Noncatalyst And Catalyst-Equipped Automobiles And Heavy-Duty Diesel Trucks. *Environmental Science & Technology* **27**, 636-651.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1993b). Sources Of Fine Organic Aerosol.3. Road Dust, Tire Debris, And Organometallic Brake Lining Dust - Roads As Sources And Sinks. *Environmental Science & Technology* **27**, 1892-1904.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1993c). Sources Of Fine Organic Aerosol.4. Particulate Abrasion Products From Leaf Surfaces Of Urban Plants. *Environmental Science & Technology* **27**, 2700-2711.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1993d). Sources Of Fine Organic Aerosol.5. Natural-Gas Home Appliances. *Environmental Science & Technology* **27**, 2736-2744.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1997a). Sources of fine organic aerosol.7. Hot asphalt roofing tar pot fumes. *Environmental Science & Technology* **31**, 2726-2730.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1997b). Sources of fine organic aerosol.8. Boilers burning No. 2 distillate fuel oil. *Environmental Science & Technology* **31**, 2731-2737.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T. (1998). Sources of fine organic aerosol. 9. Pine, oak and synthetic log combustion in residential fireplaces. *Environmental Science & Technology* **32**, 13-22.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simonelt, B. R. T. (1991). Sources Of Fine Organic Aerosol.1. Charbroilers And Meat Cooking Operations. *Environmental Science & Technology* **25**, 1112-1125.

Rumburg, B., Alldredge, R., and Claiborn, C. (2001). Statistical distributions of particulate matter and the error associated with sampling frequency. *Atmospheric Environment* **35**, 2907-2920.

Russell, A., and Dennis, R. (2000). NARSTO critical review of photochemical models and modeling. *Atmospheric Environment* **34**, 2283-2324.

Russell, A. G., McRae, G. J., and Cass, G. R. (1983). Mathematical-Modeling Of The Formation And Transport Of Ammonium-Nitrate Aerosol. *Atmospheric Environment* **17**, 949-964.

Ryan, W. M., West, C. R., Holtz, D. R., Peters, T. A., Cooper, J. A., and Ono, D. M. (1998). Reconciliation of receptor and dispersion modeling impacts of PM10 in Hayden, Arizona, In: Mathai, C. V.; Stonefield, D. H. (Eds.), Transactions, PM10: Implementation of Standards. *Air Pollution Control Association, Pittsburgh, PA.*

Saldiva, P. H. N., Pope, C. A., Schwartz, J., Dockery, D. W., Lichtenfels, A. J., Salge, J. M., Barone, I., and Bohm, G. M. (1995). Air-Pollution And Mortality In Elderly

People - A Time-Series Study In Sao-Paulo, Brazil. *Archives Of Environmental Health* **50**, 159-163.

Satheesh, S. K., and Moorthy, K. K. (2005). Radiative effects of natural aerosols: A review. *Atmospheric Environment* **39**, 2089-2110.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T. (1999a). Measurement of emissions from air pollution sources. 1. C-1 through C-29 organic compounds from meat charbroiling. *Environmental Science & Technology* **33**, 1566-1577.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T. (1999b). Measurement of emissions from air pollution sources. 2. C-1 through C-30 organic compounds from medium duty diesel trucks. *Environmental Science & Technology* **33**, 1578-1587.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T. (2001). Measurement of emissions from air pollution sources. 3. C-1-C-29 organic compounds from fireplace combustion of wood. *Environmental Science & Technology* **35**, 1716-1728.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T. (2002a). Measurement of emissions from air pollution sources. 4. C-1-C-27 organic compounds from cooking with seed oils. *Environmental Science & Technology* **36**, 567-575.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T. (2002b). Measurement of emissions from air pollution sources. 5. C-1-C-32 organic compounds from gasoline-powered motor vehicles. *Environmental Science & Technology* **36**, 1169-1180.

Schauer, J. J., Rogge, W. F., Hildemann, L. M., Mazurek, M. A., and Cass, G. R. (1996). Source apportionment of airborne particulate matter using organic compounds as tracers. *Atmospheric Environment* **30**, 3837-3855.

Schwab, J. J., Hogrefe, O., Demerjian, K. L., and Ambs, J. L. (2004). Laboratory characterizaton of modified tapered element oscillating microbalance samplers. *Journal Of The Air & Waste Management Association* **54**, 1254-1263.

Sheppard, L., Levy, D., and Checkoway, H. (2001). Correcting for the effects of location and atmospheric conditions on air pollution exposures in a case-crossover study. *Journal Of Exposure Analysis And Environmental Epidemiology* **11**, 86-96.

Singh, K. P., Warsono, Bartolucci, A. A., and Bae, B. (2001). Mathematical modeling of environmental data. *Mathematical And Computer Modelling* **33**, 793-800.

Spicer, C. W., Koetz, J. R., Keigley, G. W., Sverdrup, G. M., and Ward, G. F. (1981). A study of nitrogen oxides reactions within urban plumes transported over the ocean. *US-EPA, Research Triangle Park, NC.*

Stockwell, W. R., Middleton, P., Chang, J. S., and Tang, X. (1990). The second generation regional acid deposition model chemical mechanism for regional air quality modeling. *Journal Of Geophysical Research-Atmospheres* **95**, 16,343-16,367.

Taghavi, M., Cautenet, S., and Arteta, J. (2005). Impact of a highly detailed emission inventory on modeling accuracy. *Atmospheric Research* **74**, 65-88.

Takemura, T., Nozawa, T., Emori, S., Nakajima, T. Y., and Nakajima, T. (2005). Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model. *Journal Of Geophysical Research-Atmospheres* **110**.

Tilmes, S., and Zimmermann, J. (1998). Investigation on the spatial scales of the variability in measured near-ground ozone mixing ratios. *Geophysical Research Letters* **25**, 3827-3830.

Tobias, H. J., Beving, D. E., Ziemann, P. J., Sakurai, H., Zuk, M., McMurry, P. H., Zarling, D., Waytulonis, R., and Kittelson, D. B. (2001). Chemical analysis of diesel engine nanoparticles using a nano-DMA/thermal desorption particle beam mass spectrometer. *Environmental Science & Technology* **35**, 2233-2243.

Tonnesen, G. S. (2003). Regional haze modeling: Recent modeling results for VISTAS and WRAP. *Models-3 User's workshop, Research Triangle Park, North Carolina*.

Tsyro, S. (2002). First estimates of the effects of aerosol dynamics in the calculation of PM10 and PM2.5. *EMEP/MSC-W*, www.emep.int/reports/mscw_note_4_2002.pdf.

Turpin, B. J., and Lim, H. J. (2001). Species contributions to PM2.5 mass concentrations: Revisiting common assumptions for estimating organic mass. *Aerosol Science And Technology* **35**, 602-610.

UCAR (2003a). GCIP NCEP Eta model output. *http://dss.ucar.edu/datasets/ds609.2/*.

UCAR (2003b). NCEP ADP Global Surface Obs. *http://dss.ucar.edu/datasets/ds464.0/*.

UCAR (2003c). NCEP ADP Global Upper Air Obs. *http://dss.ucar.edu/datasets/ds353.4/*.

UCAR (2003d). NCEP ADP Global Upper Air Obs. *http://dss.ucar.edu/datasets/ds353.1/*.

UCAR (2003e). TDL U.S. and Canada Surface Hourly Observations. *http://dss.ucar.edu/datasets/ds472.0/*.

Ulrich, M. M. W., Alink, G. M., Kumarathasan, P., Vincent, R., Boere, A. J. F., and Cassee, F. R. (2002). Health effects and time course of particulate matter on the

cardiopulmonary system in rats with lung inflammation. *Journal Of Toxicology And Environmental Health-Part A* **65**, 1571-1595.

Unal, A., Odman, T., and Russell, T. (2004.). Assessing the Impacts of Hartsfield-Jackson Atlanta International Airport on Particulate Matter and Ozone in Atlanta Metropolitan Area. *Prepared For National Aeronautics and Space Administration. 2004.*

Unal, A., Tian, D., Hu, Y., and Russell, A. G. (2003). 2000 Emissions Inventory for Fall Line Air Quality Study (FAQS). *Prepared for: Georgia Department of Natural Resources, Environmental Protection Division - Air Protection Branch, Atlanta, GA.*

US-EPA (1990). CMB8 User's Mannual. *Office of Air Quality, Planning and Standards, Research Triangle Park, NC 27711*.

US-EPA (2003a). 1999 National Emission Inventory (NEI) documentation and data. *http://www.epa.gov/tth/chief/net/1999inventory.html.*

US-EPA (2003b). Clean air markets -- Data and maps. *http://cfpub.epa.gov/gdm/.*

US-EPA (2003c). Technology transfer network, clearinghouse for inventories and emission factors. *http://www.epa.gov/ttn/chief/emch/spatial/.*

US-EPA (2004a). Air Quality Criteria for Particulate Matter (October 2004). *EPA 600/P-99/002aF-bF*.

US-EPA (2004b). Biogenic emissions inventory system. *http://www.epa.gov/asmdnerl/biogen.html, ftp://ftp.epa.gov/amd/asmd/beld3, ftp://ftp.epa.gov/amd/asmd/beld3/ascii*.

US-EPA (2004c). Economic growth analysis system version 4.0 and documentation. *http://www.epa.gov/ttn/chief/emch/projection/egas40/.*

US-EPA (2004d). SMOKE User's manual. *http://cf.unc.edu/cep/empd/products/smoke/.*

US-EPA (2004e). SMOKE User's manual. *http://cf.unc.edu/cep/empd/products/smoke*.

Van Loy, M., Bahadori, T., and Wyzga, R. (2000). The Aerosol Research and Inhalation Epidemiology Study (ARIES): PM2.5 Mass and Aerosol Component Concentrations and Sampler Intercomparisons. *Journal of Air and Waster Management Association* **50**, 1446-1458.

Vautard, R., Martin, D., Beekmann, M., Drobinski, P., Friedrich, R., Jaubertie, A., Kley, D., Lattuati, M., Moral, P., Neininger, B., and Theloke, J. (2003). Paris emission inventory diagnostics from ESQUIF airborne measurements and a chemistry transport model. *Journal Of Geophysical Research-Atmospheres* **108**.

Vedal, S. (1997). Ambient particles and health: Lines that divide. *Journal Of The Air & Waste Management Association* **47**, 551-581.

Venkataraman, C., Lyons, J. M., and Friedlander, S. K. (1994). Sire Distributions Of Polycyclic Aromatic-Hydrocarbons And Elemental Carbon.1. Sampling, Measurement Methods, And Source Characterization. *Environmental Science & Technology* **28**, 555-562.

Wadden, R. A., Scheff, P. A., Lin, J., Lee, H., Keil, C., Graf-Teterycz, J., Keehan, K., Kenshki, D., Milz, S., Holsen, T. M., and Khalili, N. (1991). Two Phase Receptor Modeling. *Proceeding of the 84th annual meeting and exhibition of the Air & Waste Management Association, Vancouver, British Columbia.*

Wade, K., Marmur, A., Mulholland, J., Russell, A., Peel, J., and Klein, M. (2004). Analyses of air quality data for an emergency department study, 1993-2002. *Epidemiology* **15**, S61-S61.

Wark, K., Warner, C. F., and Davis, W. T. (1997). Air pollution. Its origin and control. *Prentice Hall*, 560p.

Watson, J. G., and Chow, J. C. (2000). Reconciling urban fugitive dust emissions inventory and ambient source contribution estimates: Summary of current knowledge and needed research. Report No. 6110.4D2. *Prepared for US Environmental Protection Agency, Research Triangle Park, NC, by Desert Research Institute, Reno, NV.*

Watson, J. G., and Chow, J. C. (2001). Source characterization of major emission sources in the Imperial and Mexicali Valleys along the US/Mexico border. *Science Of The Total Environment* **276**, 33-47.

Watson, J. G., Chow, J. C., Dubois, D., Green, M., Frank, N., and Pitchford, M. (1997). Guidance for network design and optimum site exposure for PM2.5 and PM10. *U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC.*

Watson, J. G., Chow, J. C., and Houck, J. E. (2001). PM2.5 chemical source profiles for vehicle exhaust, vegetative burning, geological material, and coal burning in Northwestern Colorado during 1995. *Chemosphere* **43**, 1141-1151.

Watson, J. G., Chow, J. C., Lowenthal, D. H., Robinson, N. F., Cahill, C. F., and Blumenthal, D. L. (2002a). Simulating changes in source profiles from coal-fired power stations: Use in chemical mass balance of PM2.5 in the Mount Zirkel Wilderness. *Energy & Fuels* **16**, 311-324.

Watson, J. G., Chow, J. C., and Pace, T. G. (2000). Fugitive dust emissions, In: Davis, W. T. (Ed.). *Air Pollution Engineering Manual, second ed. John Wiley, New York*, 117-135.

Watson, J. G., Zhu, T., Chow, J. C., Engelbrecht, J., Fujita, E. M., and Wilson, W. E. (2002b). Receptor modeling application framework for particle source apportionment. *Chemosphere* **49**, 1093-1136.

Whitby, K. T. (1978a). The physical characteristics of sulfur aerosl. *Atmospheric Environment* **12**, 135-159.

Whitby, K. T. (1978b). The physical characteristics of sulfur aerosol. *Atmospheric Environment* **12**, 135-159.

Wilemski, G. (1984). Composition Of The Critical Nucleus In Multicomponent Vapor Nucleation. *Journal Of Chemical Physics* **80**, 1370-1372.

Woo, K. S., Chen, D. R., Pui, D. Y. H., and McMurry, P. H. (2001a). Measurement of Atlanta aerosol size distributions: Observations of ultrafine particle events. *Aerosol Science And Technology* **34**, 75-87.

Woo, K. S., Chen, D. R., Pui, D. Y. H., and Wilson, W. E. (2001b). Use of continuous measurements of integral aerosol parameters to estimate particle surface area. *Aerosol Science And Technology* **34**, 57-65.

Yan, B., Zheng, M., and Russell, A. (2004). Uncertainty and Sensitivity Analysis of Chemical Mass Balance Modeling Using Organic Tracers for $PM_{2.5}$ Source Apportionments. *Proceeding of the AAAR 04 annual conference, Atlanta, GA.*

Yang, Y. J., Wilkinson, J. G., and Russell, A. G. (1997). Fast, direct sensitivity analysis of multidimensional photochemical models. *Environmental Science & Technology* **31**, 2859-2868.

Ying, Q., Mysliwiec, M., and Kleeman, M. J. (2004). Source apportionment of visibility impairment using a three-dimensional source-oriented air quality model. *Environmental Science & Technology* **38**, 1089-1101.

Zeng, Y., and Hopke, P. K. (1989). A Study Of The Sources Of Acid Precipitation In Ontario, Canada. *Atmospheric Environment* **23**, 1499-1509.

Zhang, Q., Canagaratna, M. R., Jayne, J. T., Worsnop, D. R., and Jimenez, J. L. (2005). Time- and size-resolved chemical composition of submicron particles in Pittsburgh: Implications for aerosol sources and processes. *Journal Of Geophysical Research-Atmospheres* **110**.

Zhang, Y., Pun, B., Vijayaraghavan, K., Wu, S. Y., Seigneur, C., Pandis, S. N., Jacobson, M. Z., Nenes, A., and Seinfeld, J. H. (2004). Development and application of the model of aerosol dynamics, reaction, ionization, and dissolution (MADRID). *Journal Of Geophysical Research-Atmospheres* **109**.

Zheng, M., Cass, G. R., Schauer, J. J., and Edgerton, E. S. (2002). Source apportionment of PM2.5 in the southeastern United States using solvent-extractable organic compounds as tracers. *Environmental Science & Technology* **36**, 2361-2371.

Zheng, M., Ke, L., Wang, F., Cass, G. R., Schauer, J. J., Edgerton, E. S., and Russell, A. G. (2005). Source apportionment of PM2.5 and their daily variations at Jefferson Street, Atlanta GA during Summer and Winter. *Journal of Air and Waste Association. Submitted for publication*.

VITA

SUN-KYOUNG (HELENA) PARK

SUN-KYOUNG (HELENA) PARK received a B.S. in Earth Science Education in 1997 and an M.S. in Atmospheric Science in 1999 from Seoul National University in Korea. She received an M.S. in Environmental Engineering in 2000 from Georgia Tech, where she worked as a graduate research assistant pursuing a Ph. D. degree concentrated on Environmental Engineering since 1999. Her main research areas are air quality modeling and receptor modeling. Main applications include source apportionment of $PM_{2.5}$ and epidemiological studies of air pollution.